The Open University

S103 discovering science

Articles for Block 12: Part 1

Contents

(For Articles 5, 7 and 10 see Articles for Block 12: Part 2)

1.5 DS/P

SUP 43585 7

The age of life on Earth

The age of the Earth is now reliably set at 4.5 billion years. Up through the 1950s people thought that life as we know it probably did not appear until recently—about one billion (10^9) years ago. This was largely because fossils had not been found in rocks older than 600 million years. (It also reflects a prejudice that life's origin represents a series of highly unlikely events, and so took a "long time" to occur.) However, sedimentary rock sections viewed under the microscope often reveal fossil microorganisms, and these tell an entirely different story. Some sediments one billion years old are teeming with fossil algae—recognizably like certain existing types. At two billion years, sediments from the Lake Superior region show a variety of unusual microscopic forms. This is about the time geologists estimate that oxygen reached appreciable levels in the atmosphere. By three billion years ago, the variety is no longer there, but a few rocks this old contain simple bacteria-like fossils, some apparently even in states of division. The oldest known microfossils, in rocks from the "North Pole" area in the north of Western Australia, are 3.5 billion years old (Fig. 19). All fossils in these ancient rocks (two billion years and older) appear to be bacteria. Evidence for eukaryotes at these times has yet to be found. Most people argue that eukaryotes did not arise until oxygen appeared in the atmosphere and an ozone shield developed around the atmosphere, cutting down the influx of destructive ultraviolet light. Geological estimates indicate that O_2 levels in the Earth's atmosphere became appreciable 2.0 to 1.5 billion years ago.

Another important type of fossil evidence is a macroscopic structure called the stromatolite. Photosynthetic cyanobacteria (blue-green algae) growing in concert with other bacteria form large, mat-like structures resulting from the deposition of minerals in the colonial bacterial growth. Although most *stromatolites*, as these are called, are produced by cyanobacteria, they can also be produced by another photosynthetic bacterium, *Chlorofexus*. Stromatolites have a characteristic striated structure (Fig. 20) and are common features in many sedimentary rocks throughout the ages, even those of 3 and 3.5 billion years ago. They tell us that photosynthetic bacterial life almost certainly existed as early as 3.5 billion years ago.

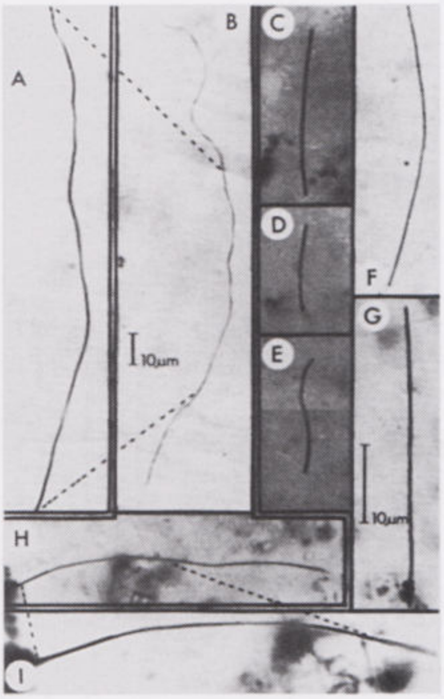


Figure 19. Filamentous microfossils in black chert from the Warrawoona Group of the North Pole Dome area of Western Australia (about 1250 km NNE of Perth). Age 3.5 billion years. (Courtesy J. W. Schopf.)

Figure 20. Stromatolites. Two types of photosynthetic bacteria, the blue-green or cyanobacteria and *Chloroflexus*, form colonial masses in which they are the dominant organisms. These formations, called stromatolites, may be small or large. Minerals deposited in stromatolites as they form make them readily fossilizable. Stromatolites have a characteristic laminar structure reflecting seasonal variation during their formation, which makes them easy to recognize in sedimentary rock formations, where they frequently occur. Their presence in 3.5 billion-year-old rocks indicates that photosynthetic bacteria—presumably of the *Chloroflexus* or cyanobacterial variety—had evolved by that time. (Courtesy E. S. Barghoorn.)

The fossil record, unfortunately, cannot tell us how far back in time bacterial life extends. The 3.5-billion-year figure is the end of the line. No older well-preserved sedimentary rocks are yet known. Rocks from Greenland's Isua formation are 3.8 billion years old. These were sedimentary rocks to begin with, but they have unfortunately suffered so much heating and compression that much of their original structure has been altered. Microfossils tend to be destroyed by this process. Even so, these rocks still retain evidence for the possible existence of life. Living systems distinguish ever so slightly among the *isotopes* of carbon and other elements. For example, when carbon has passed through a chain of bacteria, the ratio of C_{12} to C_{13} is increased. This permits sedimentary carbon derived from living systems to be distinguished from carbon of non-living origin. Carbon-rich areas in the Isua rocks clearly show isotope ratios characteristic of living systems.

The fossil record tells us two important things: life existed at least as far back as 3.8 billion years ago, and by 3.5 billion years ago some sort of photosynthetic bacteria had evolved.

Genetic evidence

How much further back in time does life extend? To try to answer this we have to turn to another type of record—that written into the cell's genes. To see how genetic sequence can serve as a time measure, picture a castaway on a deserted island marking time by carving notches in a stick. The original condition of the stick is unnotched. Each notch added represents some time interval, a day, an hour, a season. By summing the notches the castaway can tell how long he has been on the island. A genetic sequence existing at some time in the past corresponds to the unnotched stick. As time goes by, changes (due to gene mutations) occur in the sequence; these are the notches. By counting the number of changes, we can find the number of intervals that have gone by. Unfortunately, we do not know what the original sequence was, only the present version. Therefore, we resort to a trick. We compare two different current versions of the sequence (taken from two different organisms). Where the two are identical, we assume they are the same as their common ancestral version. Where they differ, one of them must have changed in the intervening time since they shared a common ancestor; that is, one of them has become "notched." (As the reader may see, this analogy is simplistic and so only approximately true.) Therefore, the number of differences between the two indicates how long—on a relative time scale—it has been since the time of a common ancestor for the two sequences or organisms. In principle, such an approach both measures relative time intervals and determines genealogies.

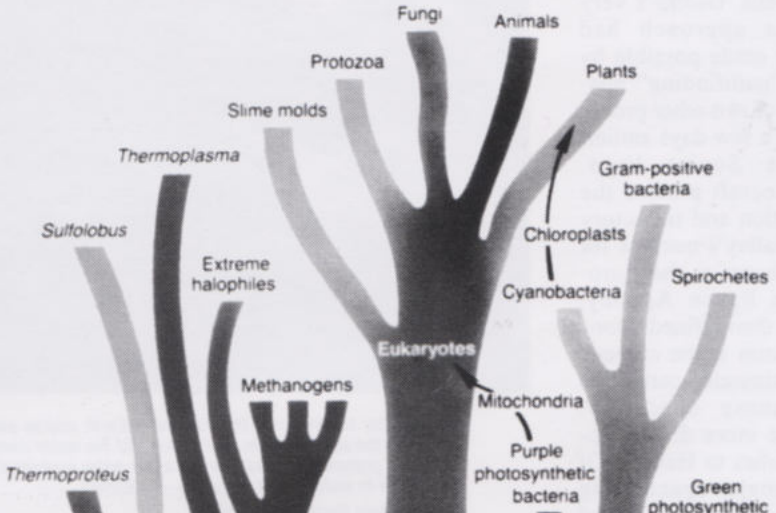
The biologist now has at his command the technical capacity to sequence any isolatable gene or gene product, be it RNA or protein. He is in a position to do genealogical measurements of the sort described.

The most useful system to date is based on comparative analysis of ribosomal RNA sequences. By this method a genealogical tree that covers all existing life can be constructed. Vestiges of the ribosomal RNA sequence of the ancestor common to all present life still remain! The universal genealogical tree is shown in Figure 21. The common ancestor of all has given rise to three main branches. The first branch is the *true bacteria* (eubacteria). These are typical bacteria familiar to any student of biology. The cyanobacteria and other photosynthetic bacteria are included in this group. The second branch is the *archaebacteria*. These are a little-known group of prokaryotes that includes the *methanogens*, a type of organism that may figure prominently in our considerations. Archaebacteria are not specifically related to true bacteria. Both are just as closely related to the remaining major branch. This third branch is the *eukaryotic cell*, or at least some major aspect of it. It may represent the "host" organism discussed above in connection with the evolution of eukaryotic organelles.

How old is the universal common ancestor? If photosynthetic bacteria became established by 3.5 billion years ago, then the common ancestor of eubacteria—the group to which the photosynthetic forms belong—must have existed at a still earlier time. But even this follows the era when the *universal* common ancestor existed. Obviously, times cannot be assigned with certainty, but it seems reasonable that the earliest detectable life form, the universal common ancestor, existed at least four billion years ago. In a sense, a pendulum has come full swing. Thirty years ago, we felt life had arisen relatively recently—one billion years ago—and wondered why it took so long. Now we see it arising so near to the time Earth was formed that we must wonder how this possibly could have happened. What the fossil record and the genealogical evidence are telling us is that the very early events in the geologic history of the planet may also be the crucial events in the origin of life.

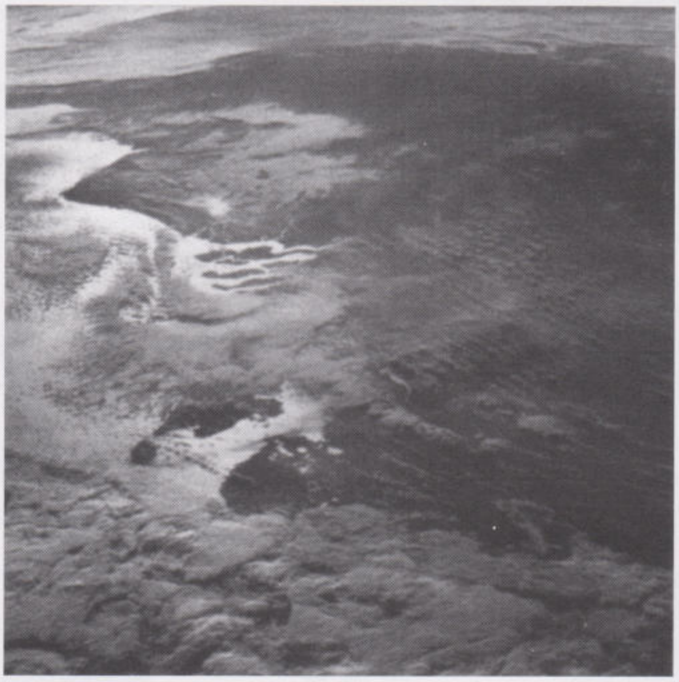
Figure 21. The universal phylogenetic tree. All known living forms fall into one of the three basic categories: the archaebacteria, the eubacteria, and the eukaryotes. The first two are prokaryotic cells; that is, they are small, do not have nuclear membranes or organelles, etc. However, the archaebacteria and eubacteria are not specifically related to one another. From fossil evidence bacteria appear to have been in existence for at least 3.5 billion years and perhaps longer, far before the advent of the eukaryotic groups. Each of the three major groups of organisms is quite distinct on the molecular level, differing from the others in the molecular details of their common cellular processes.

[For a colour version of this figure see *Articles for Block 12: Part 1*.]



Seeding Earth: Comets, Oceans and Life

by Christopher Chyba



Liquid water flowing over the surface of Earth makes our planet unique in the solar system. But where did the water come from? Did Earth process it herself, or did other solar system objects contribute to water—and so to life—on Earth?

Photo: Johnson Space Center/NASA

Giotto, the European spacecraft sent to encounter Halley's Comet in March 1986, pierced the cloud of gas and dust surrounding the comet's core and flew within 600 kilometers (400 miles) of its frozen nucleus. *Giotto*'s very close approach had been made possible by the "pathfinding" mission of two other probes only a few days earlier; these Soviet *Vega* spacecraft plotted the position and trajectory of Halley's nucleus for engineers at the European Space Agency, who then refined *Giotto*'s aim at the comet's gas-shrouded core. Two Japanese spacecraft made more distant approaches to Halley. (Of the major spacefaring nations, only the United States sent no representative to comet Halley. However, NASA's Deep Space Network of radio telescopes did make an essential contribution to the precise tracking required for the Soviet pathfinding mission.)

Observations by the spacecraft, complemented by observations made with instruments on the ground, in sounding rockets and in high-altitude aircraft, verified the long-standing hypothesis that comets consist largely of frozen water: Halley is about half water ice.

The suspicion that comets are loaded with organic molecules, the kinds of molecules needed for the earliest stages in the evolution of life on Earth, was also dramatically confirmed: The *Giotto* and *Vega* spacecraft sampled the comet's dust as they flew through its gaseous envelope and found organic-rich microscopic grains. (In the language of chemistry, saying a molecule is "organic" means only that it is a carbon compound, often with hydrogen, nitrogen or other atoms—the term carries no biological connotation.)

Some of the cometary grains, altogether free of rock, seemed to be purely organic in composition. But Halley's solid center held the big surprise: It was much larger, and much blacker, than many investigators had expected. We now think that the carbon-black exterior of the nucleus may be a kind of webbed crust of organic molecules and rock fragments.

If Halley Hit a Planet

Ground-based observations suggest that other comets are like Halley in gross respects; for the time being the best we can do is to assume that they are like Halley in detail as well. Suppose, then, that most comets are 50 percent water ice. Suppose that most comets are full of organic molecules. What would be the result of a cometary collision with a planet?

Comets course through the solar system in elongated orbits. Halley, for example, hurtles out past the orbit of Neptune and then, on its return, reaches the orbit of Venus, crossing the path of every planet in between. Sooner or later it will have a close encounter with one of these planets. Such an encounter might deflect the comet to a new orbit, perhaps ejecting it from the solar system altogether. But if the encounter is close enough, the planet's gravity will pull the comet into a collision.

Striking the surface of the planet, the comet would excavate a crater and vaporize almost entirely in the ensuing explosion. The comet's steam might form a tenuous atmosphere of water vapor, or perhaps clouds if the planet already has an atmosphere. Some of the comet's organic molecules might survive the explosion. In sum, the impact of a Halley-type comet would deliver to the surface of the target world a certain amount of organics and a great deal of water—around 250 cubic kilo-

meters' worth.

HALLEY'S COMET OVER ANTARCTICA—*The water that covers Earth today may have originated in comets that condensed in the outer regions of our solar system.* Painting: Kim Poor

The most recent cometary collision with Earth probably took place in 1908 when a huge explosion over Tunguska, Siberia caused widespread devastation in what was fortunately an extremely remote area. Throughout its history, Earth must have experienced many such collisions. But how many? How much cometary water might Earth have collected in this way?

Different theoretical models give different answers, but we can place empirical limits on these answers by using knowledge gained from the *Apollo* missions to the Moon and probes to other planets.

Early Bombardment

When *Apollo 11* astronauts returned to Earth in July 1969, the first lunar samples quickly established two facts of enormous importance for the Moon's history: Most of the surface of the Moon is very ancient, and most of this surface has been geologically inactive for the past 3 billion years. Over the following three years of *Apollo* landings, it became clear that the oldest lunar terrain shows the cratering record from a period of intense bombardment that ended about 3.8 billion years ago.

What projectiles collided with the Moon to excavate these craters, and did the same family of objects also batter Earth? No terrestrial evidence can answer these questions. The oldest known rocks on Earth, from the Isua formation in Greenland, are no more than 3.8 billion years old. (However, in October 1989 Samuel Bowring at Washington University in St. Louis announced finding two rocks 3.96 billion years old in remote northern Canada.) Earth is such a geologically active planet that whatever terrestrial surface existed that long ago has since been completely eroded away or drawn back inside Earth's hot interior to be melted and forever lost as geological evidence. However, *Mariner* photographs of Mars and Mercury from the early 1970s demonstrate that the family of objects responsible for the lunar cratering bombarded these planets as well.

"Size/frequency" graphs, reflecting the number of craters as a function of crater size for each body, show definitive similarities in the impact history of Mercury, the most ancient terrain on Mars and the most heavily cratered terrain on the Moon. All these worlds, and by implication every planet in the inner solar system, were bombarded by the same family of projectiles, identifiable by their characteristic size/frequency distribution. Although direct evidence has been removed from its surface, Earth must have experienced this intense bombardment as well, as objects with diameters up to hundreds of kilometers repeatedly traversed the region of the

THE ICEBERGS—*Only on Earth are temperatures near water's triple point, where it can exist simultaneously as vapor, liquid and ice, as seen in this dramatic rendering inspired by the artist's expedition to Labrador.*
Painting: Frederick E. Church, courtesy Dallas Museum of Art

inner planets. Mercury, Venus, the Earth-Moon system and Mars gradually "swept up" these rogue bodies in collision after collision.

Comet-Delivered Oceans

In a recent study, I used the cratering record on the Moon to estimate how many of these bodies struck the lunar surface during the period of intense bombardment. From that number, taking into account Earth's larger diameter and higher gravity, I calculated the total mass of impactors collected by Earth during that time. Between 4.5 and 3.8 billion years ago, it appears that Earth swept up a total mass as great as a third of the mass of the Moon.

It is not certain how many of these wayward bodies were comets. Even if only a small fraction were comets, the implications for the early Earth are profound. A cometary fraction of about 10 percent would have been sufficient to deliver an ocean's worth of water to Earth's surface.

We don't have clear evidence to determine whether comets comprised 10 percent of the impacting bodies during the intense bombardment. In current cratering in the inner solar system, comets appear to account for between one tenth and one third of the total, although some investigators believe the fraction may actually be as high as one half. But cometary cratering in the present tells us little about its importance in the past.

Dating the Early Bombardment

Our knowledge of the early heavy bombardment of the inner solar system derives almost entirely from examination of the samples brought back from the Moon by the *Apollo* astronauts and by Soviet *Luna* robotic missions. From these samples we discovered that lunar cratering was not uniformly distributed through time; rather it decreased exponentially until about 3.8 billion years ago, when the frequency of cratering leveled off to its present low rate. Thus the ancient environment of all the planets of the inner solar system—including Earth's at the time of the origin of life—was impact-dominated, with cratering thousands of times more intense than today.

This was indeed a great deal to learn from the painstaking research lavished on the 380 kilograms (840 pounds) of rocks returned from the Moon. How was it accomplished?

The first step was to determine the age of the samples. While a number of radioactive dating methods are in use, the fundamental idea is well illustrated by a simplified version of the potassium-argon method.

Argon is a "noble" atom, one that does not chemically combine with other atoms. Thus when a rock "contains" argon it literally encloses this gaseous element: The argon is physically prevented from leaving the sample by the layers of rock surrounding it. If the rock is liquefied, the argon will bubble out and escape.

The Sea of Tranquillity, where *Apollo 11* landed, is a vast lunar lava flow, and when this flow was molten, it must have lost whatever argon it contained. But the rock samples returned by the astronauts contained traces of the isotope argon 40. This isotope is produced in the radioactive decay of an unstable isotope, potassium 40, whose decay rate can be precisely measured in the laboratory.

We add the amount of potassium 40 now remaining plus the amount of argon 40 decay product, and this gives us the initial amount of potassium present when the lava solidified. The decay rate then tells us how long it must have taken this much initial potassium 40 to transform into that much final argon 40. That length of time is the age of the rock.

Once we know the age of a particular locale, we can use "crater counts" to determine the relative age of other sites (more heavily cratered locales are older than less cratered ones). A knowledge of the ages of many different sites then allows us to determine the number of lunar craters formed in any given interval of time.

Thus we owe our understanding of the early bombardment of the inner solar system to a few Moon rocks, made available for study by sample-return missions. Similar missions to Mars will test these conclusions and undoubtedly provide new surprises. —*CC*

If comets brought a significant quantity of water to Earth, we would expect to find that they delivered water to the other planets as well. But again the evidence is indecisive. Mercury and the Moon appear to have no water, but water would not last long on either of these bodies, except perhaps in rare permanently shadowed regions. On the other hand there is indirect (though controversial) evidence for much water on Venus in the distant past (see the November/December 1988 *Planetary Report*). As for Mars, *Mariner* and *Viking* photographs show a variety of now-dry water channels and traces of ancient lakes. Early Mars may have been warmer than at present, with water flowing on its surface. So far there is no way to test whether this water was delivered by comets.

Closer to home, we can readily test Earth's ocean water, the isotopic composition of which hints at an extraterrestrial origin. In typical ocean water only about one water molecule in ten thousand contains deuterium (D), a hydrogen isotope that carries a neutron and is thus heavier than normal hydrogen (H). The D/H ratio in Earth's oceans is higher than the "cosmic" D/H ratio typical in our solar system and elsewhere in the galaxy by a factor of about ten.

Observations of Halley indicate that the D/H ratio of cometary ice is about equal to that of Earth's oceans—elevated above the cosmic D/H value by about the same factor of ten. Moreover, volcanic water from deep within Earth's interior appears to be closer to the cosmic D/H ratio; this evidence does not confirm but is consistent with an explanation of Earth's surface water as a late-arriving cometary veneer.

An important theoretical objection calls into question a cometary origin for Earth's oceans. Recent work by Jay Melosh and Ann Vickery of the University of Arizona shows that some impacts of comets and asteroids may erode planetary atmospheres by blasting atmospheric gases off into space.

If this view is correct, then cometary impacts would not simply have delivered water to Earth but would also have eroded it away. Research I have recently completed, however, seems to show that for the larger worlds of the inner solar system the competition between these two processes strongly favored the net accumulation of planetary oceans. This same work also points to the possibly important role of water-rich carbonaceous asteroids. If Earth's water arrived as a late veneer, it is like-

ly that asteroids and comets both were important in its delivery.

Seeding Earth with Organics

If comets made a significant contribution to the oceans of primitive Earth —the conclusion remains tentative— then they played a major role in shaping the environment in which life evolved. Comets may also have contributed to the terrestrial inventory of organic molecules necessary for the origins of life.

In 1953 Stanley Miller, a University of Chicago graduate student working with the Nobel Prize-winning chemist Harold Urey, showed that amino acids and other organic molecules would form easily and naturally in what was then believed to have been the atmosphere of the early Earth. Simulating this primitive atmosphere with a gas mixture of methane, ammonia and water, Miller introduced an electrical spark (representing, for example, lightning) and obtained a high yield of amino acids.

It has since been shown that in such an atmosphere virtually any energy input (for example, ultraviolet light from the Sun, or shock energy from meteoritic impacts) will lead to the creation not only of amino acids but to the precursors of other important biological molecules. In the traditional view of primitive Earth, these precursor molecules collected in the oceans, forming a warm, dilute "organic soup," on the surface or shorelines of which life evolved.

More recently consensus among geochemists has been shifting to the view that the early terrestrial atmosphere may not have been so biologically accommodating. In this picture, the early atmosphere consisted not of methane and ammonia but mainly of nitrogen and great quantities of carbon dioxide. Under these conditions production of organic molecules is more difficult, and thus environments suitable for the origin of life on primitive Earth would have been more rare. Whichever scenario is correct, comets may have provided an important extraterrestrial link for the origins of terrestrial life.

Giotto and *Vega* showed that comets are full of organic molecules. It now appears that such molecules form inevitably in comets by a process analogous to the Miller/Urey experiment. Miller and Urey used a gaseous mixture, but the experiment will also produce organics when methane, ammonia and water are present in solid form, as ices.

(continued on page 30)

TOP: From Earth orbit we gain a perspective on water molecules cycling through their phases that is impossible from Earth. In this photograph taken by Apollo 9 *astronauts, we look almost directly down on a thunderhead forming over South America.*

MIDDLE: The perspective from Earth orbit also enables us to see the extent of humanity's effect on our planet. Here Challenger *astronauts looked down on growing cumulonimbus clouds over Zaire, but haze created by fires set to clear ground for agriculture completely obscured the surface.*

BOTTOM: Icebergs large enough to be seen from space break from an ice floe and drift off the coast of Labrador.

Photographs: Johnson Space Center/NASA

Seeding Earth

(continued from page 23

Both before and after cometary formation, the icy grains that make up comets are subjected to many kinds of radiation. Experimenters in North America and Europe have irradiated likely recipes of cometary ices in the laboratory, and the result is inevitably the formation of organic molecules. Recent work I have done in collaboration with Carl Sagan of Cornell University shows that certain spectral features observed in comet Halley, as well as in the lesser-known comets Wilson and Bradfield, are probably due to such irradiation-derived organics.

Cometary collisions with Earth would necessarily contribute some of these organics to Earth's prebiotic inventory. However, we don't know what fraction of cometary organics would survive the high temperatures and pressures associated with the resulting explosion and crater excavation. Cometary water ice would evaporate, enter the atmosphere as steam and eventually rain out, but the cometary organics might well be destroyed. The fact that some meteorite fragments contain intact organic molecules proves that some organics can survive such collisions. Many scientists are working to determine quantitatively the survival of comet-delivered organics under plausible early-Earth conditions.

Cometary Devastation

The intense bombardment of early Earth may have resulted in an "impact frustration" of the origins of life. Large impacts would have been catastrophic for local environments, and the extreme temperatures generated by the violent collisions may have effectively sterilized vast expanses. Moreover, the huge explosions would have vaulted enormous quantities of dust into the atmosphere. There may even have been enough debris to envelop Earth in a dust cloud, blocking sunlight and creating conditions like those envisioned in the "nuclear winter" hypothesis proposed by David Grinspoon and Carl Sagan. Such an era of inhospitable conditions would mean that the time available on the early Earth for the origin of life was even shorter than previously believed.

If the terrestrial evolution of life did occur very rapidly, then the possibility increases that life may have arisen on Mars during its apparently brief, comparatively Earthlike youth. Indeed, one of the objectives of Soviet and US missions to Mars in the next decade is to look for fossil evidence of extinct martian microorganisms. The new field of "exopaleontology" will begin in the 1990s.

The oceans have long inspired feelings of wonder. It is as though an inchoate understanding of our origins visited our minds even before the theory of evolution taught us that our beginnings lay in the sea. A special tinge of awe comes with the thought that oceans—this most terrestrial image, the cradle of life—may in fact have had an extraterrestrial origin. In this realization, scientific and aesthetic appreciation join in a single moment. Many more such moments await our continued exploration of space.

Christopher Chyba studied theoretical physics and the history and philosophy of science at the University of Cambridge; he is now a graduate student in space sciences at Cornell.

The SCIENCES
MAY–JUNE 1995

Article 3

COLD START

The first life may have evolved out of a chilled organic soup that brewed under a thick layer of ice

BY JEFFREY L. BADA

SOME STORIES ARE SO COMPELLING THEY bear retelling many times. So it is with scientific stories about the beginnings of life. Forty-two years ago, at the University of Chicago, a young graduate student named Stanley L. Miller set out to reproduce the conditions that gave rise to life on earth. With a simple circuit of glass tubing he connected two flasks. Into one flask he poured water, an early ocean in miniature: into the other he released a mixture of methane, ammonia, and hydrogen, to represent primeval air. He zapped the atmosphere nonstop with an electric arc, the laboratory equivalent of lightning. When he turned off the spark a week later, the miniature ocean had turned the color of tea. Chemical analysis showed that it was loaded with organic compounds, including several amino acids, the building blocks of proteins—key players in the processes of life.

The scientific community was jubilant. "If God didn't do it this way," Miller's adviser, the Nobel laureate chemist Harold C. Urey, remarked half jokingly, "he missed a good bet." The bet he had in mind was an atmosphere with a strong tendency to trigger chemical reactions by donating electrons—in chemical terms, a reducing environment. Later, Miller as well as other scientists, including the planetary astronomer Carl Sagan of Cornell University, ran similar experiments with various reducing mixtures of gases and with other sources of energy, such as ultraviolet light. The almost invariable result: a rich organic brew—prebiotic soup.

Then came the backlash. Biochemists showed that to turn even the richest prebiotic soup into a single-celled organism would take more than time and good wishes. Some even suspected that proteins, and thus amino acids, had played little or no role in the earliest organisms. Atmospheric chemists, meanwhile, had found what they considered fatal flaws in the mixtures of gases that Miller and his successors had used in their reactions. In the real atmosphere, they pointed out, methane and ammonia would have quickly disintegrated under the influence of ultraviolet light. Most came to favour a more stable mixture of gases with drastically different chemical properties.

In recent years many investigators studying the origin of life have taken an approach different from the one Miller took: they have worked backward in time to determine how the intricate machinery of the living cell might have arisen from simpler precursors. Most workers now believe that some of the most fundamental features of modern cells (including the genetic code based on DNA and a biochemistry dependent on enzymes made of proteins) did not exist in the earliest life. Many biochemists now speak of an RNA world—a realm of life in which, before the first modern cells emerged 3.5 billion years ago, all of the major functions of life were carried out by ribonucleic acids, chemical cousins of DNA. [See Christian de Duve's article "Prelude to a Cell", *The Sciences,* November/December 1990.] Some investigators are trying to trace life back further still, into even more exotic pre-RNA worlds in which the processes of life were conducted by a much different set of organic polymers.

Stanley Miller's back-to-basics approach is still thriving, however. In the NASA Specialized Center of Research and Training in Exobiology, Miller and I and colleagues at the University of California, San Diego, and the Scripps Research Institute and the Salk Institute for Biological Studies in La Jolla, are investigating the conditions that gave rise to life on earth. We start with a simple premise: that at least 3.8 billion years ago organic molecules in the earth's ocean initiated the self-replicating chain reaction that can be regarded as the beginning of life. Then we ask: Where did the prebiotic soup come from? How could the earth and its atmosphere and oceans have developed into a reactor capable of producing it? What sort of reactor would have brewed the *best* prebiotic soup? What other steps would have been necessary to set the stage for the origin of life?

Our work so far has thrown out some clear challenges to the conventional wisdom. Among other surprising conclusions, Miller and I believe that life on earth started in what at first glance looks like a most unlikely place: in a frigid ocean, under hundreds of feet of ice.

IN MANY WAYS, THE EARTH ON WHICH LIFE emerged 3.8 billion years ago was an alien world. It was a young planet, having condensed from dust and gas only 700 million years earlier. Its surface was probably about 97 percent water. There was no oxygen in its atmosphere. Its interior was more radioactive, and thus hotter, than it is today. The sun, however, was much cooler, shining (according to the latest models of stellar evolution) 20 to 30 percent less brightly.

The early earth was violent, subject to frequent collisions with comets, meteorites, asteroids and other debris left over from the formation of the solar system. According to a scenario known as the impact-frustration hypothesis, the bombardment was fierce enough at first to boil the seas, killing off any early organisms: only about four billion years ago did things settle down. Then life developed swiftly and never died out.

All of those characteristics place important constraints on chemical models of the early earth. The persistence of life is particularly important: it implies that for almost four billion years the overall temperature of the earth has stayed between zero and one hundred degrees Celsius, the freezing and boiling points of water.

The lower end of that temperature range is particularly dangerous, because the earth, once frozen, would be hard to thaw. Ice is more than three times as reflective as water (in technical terms, it has a higher albedo), and once it forms, it repels heat that otherwise would melt it. If the earth were to freeze tomorrow, it would take a star 30 percent more luminous than the sun to melt it. For that rea-

An Early Earth?

Jupiter's moon Europa could be a good model of what our planet looked like four billion years ago

"The past is a foreign country," the English writer L. P. Hartley observed in his novel *The Go-Between*. For Jeffrey Bada and other proponents of an ice-covered early earth, however, the past might well be orbiting around another planet. Searching for evidence to bolster their theories, they have combed the solar system for worlds that might shed light on the geophysical and geochemical processes that once held sway closer to home. One of the most promising candidates is a frigid body slightly smaller than the earth's moon and five times as far from the sun as the earth is. It is Europa, the next-to-innermost large moon of the planet Jupiter.

In February, Europa made headlines when astronomers studying images from the Hubble Space Telescope concluded that oxygen is present in its atmosphere. The oxygen is a byproduct of water vapor rising from the material on Europa's surface: a thick shell of slightly dirty, rock-hard H_2O—water ice. It is the ice, not the oxygen, that excites Bada and other investigators. As Europa orbits Jupiter, tidal stresses from the giant planet's varying gravitational pull squeeze the moon, causing internal friction. In the early 1980s the planetary scientist Ray T. Reynolds of the NASA Ames Research Laboratory in Moffett Field, California, and other investigators showed that, in theory, the friction should generate enough heat to maintain a layer of liquid water beneath the surface. The result could be a global ocean capped with a frozen ceiling kilometers or even tens of kilometers thick.

In short, Europa might resemble a scaled-down working model of Bada and Miller's frozen earth. That could be a godsend for theorists trying to trace the evolution of our planets—if the hidden ocean is really there. Is it? Nobody can be sure, Reynolds cautions, but there is some tantalizing circumstantial evidence. Images transmitted back to earth in the 1970s by Pioneer and Voyager space probes revealed that the icy surface of the Jovian moon is scored with dark lines that look like fissures. Apart from that, however, Europa is amazingly smooth—a sharp contrast to the meteor-pocked exteriors of its neighboring moons Ganymede and Callisto.

"There should be craters of all ages on the surface, and we don't observe them." Reynolds says. "What that means is that something is removing them."

That something could be a sign of a sub-Europan ocean. If liquid water does flow beneath the surface, then, Reynolds suspects, the relatively warm ice above it would tend to creep, gradually erasing large craters. Small craters could perish more dramatically, when fractures in the ice reach the ocean, unleashing geysers of cold steam that recondense in a rain of frost. In principle, such geysers could be detected from the earth by a sharp rise in the water content of Europa's atmosphere. No one has ever seen anything like that, but Reynolds says he will be looking hard at the high-resolution images the *Galileo* space probe will send back when it reaches the Jovian system in early December.

All of which is a far cry from saying that Europa does, or could, harbor life. On the contrary, says Doyle T. Hall, an atmospheric scientist at Johns Hopkins University and a member of the team that discovered the oxygen: every part of Europa is hostile if not inimical to life. The atmosphere is negligible. The ocean, if one exists, is swathed in perpetual darkness and burdened with immense pressure from the overlying ice. And in between, on the moon's surface, lies a radioactive hell.

"Europa is orbiting inside Jupiter's magnetic field, and inside that magnetic field there are a lot of energetic particles—radiation to you and me," Hall says. "If you could magically go there and stand on Europa, it would probably be like getting a chest X ray every second or two. Basically, any organic matter on the surface is going to be torn apart."

That's a pity, because in some respects Europa has the makings of an attractive piece of real estate. Five billion years from now, Reynolds says, the sun will enter the red-giant phase of stellar evolution. As it swells to engulf Mercury and Venus, its heat will turn the earth into a cinder and, somewhat later, melt all the ice on Europa. "It won't last very long, but it looks as if there will be a period of a few hundred million years, very far in the future, when there may be oceans on Europa."

—Pamela Sue Frost

son most theorists take it as axiomatic that the earth has never frozen, and models of atmospheric evolution almost always include a generous dollop of greenhouse gases—especially in the days of the faint young sun.

Both methane and ammonia are greenhouse gases, and in theory, a Miller–Urey atmosphere could trap enough heat to keep the earth from freezing. The problem is that each gas would quickly break down under sunlight. That is why most atmospheric scientists now favour an early atmosphere of nitrogen and carbon dioxide, similar to the atmosphere today but with two major differences: it was free of oxygen, a later by-product of life, and it included thousands or even tens of thousands times as much carbon dioxide, necessary to offset the lower solar luminosity.

The only thing wrong with the carbon dioxide–nitrogen atmosphere is that it is useless for engendering life. Chemically it is all but inert. Pump it into Stanley Miller's flasks and jolt it with electricity, an experiment that Miller and others have tried, and you will get only carbon dioxide and nitrogen—nothing more exciting than that.

So where on earth did prebiotic soup come from?

The answer, according to some biochemists and planetary scientists, is that it came from *nowhere* on earth. Conditions on the early earth were too hostile, they say: life or the chemical precursors of life, must have come from outer space. That idea surfaced early in this century, when the Swedish physicist and chemist, Svante A. Arrhenius, proclaimed with almost missionary zeal a hypothesis he called panspermia (seeds everywhere): life, he said, had drifted to earth in a cloud of space-faring spores or micro-organisms. In 1981 the English molecular biologist and codiscoverer of the structure of DNA, Francis H. C. Crick, proposed an even more radical hypothesis: directed panspermia, the idea that the seeds of life were sent to earth by intelligent extraterrestrials. Such ideas are impossible to test, and they beg the real question: How did the supposed seeding organisms, intelligent or otherwise, come to exist?

Other investigators have suggested that what came from space was not life itself but the raw ingredients of life, organic molecules. That possibility is worth investigating. Simple organic molecules such as hydrogen cyanide (HCN) and even ethyl alcohol (C_2H_5OH) have been detected spectroscopically in interstellar dust clouds, and more complicated organic molecules—including amino acids—have been found in the kind of meteorites known as carbonaceous chondrites. The Murchison meteorite, which fell near a small town in Australia in 1969, contains at least seventy-four amino acids, including eight or more of the twenty amino acids that make up proteins in organisms on earth.

Did extraterrestrial molecules stock the prebiotic soup? To test that possibility my colleagues and I have investigated whether extraterrestrial molecules are still falling to earth

or have done so in the geologically recent past. As an indicator of such molecules we tested for alpha-aminoisobutyric acid (AIB), an amino acid that is extremely rare on the earth but common in some extraterrestrial bodies that have fallen to earth. We searched for this amino acid in ice cores from Antarctica and Greenland and in fallout from two major impacts: the 1908 Tunguska event, in which an object from space exploded over Siberia with the force of a small nuclear blast; and the famous sixty-five-million-year-old impact at the Cretaceous–Tertiary (KT) boundary, which may have caused the extinction of the dinosaurs.

Unfortunately for the theory of extraterrestrial seeding, in all the samples we checked the amounts of AIB deposited were either undetectable or pitifully small. Only one ice sample, approximately 4,500 years old, showed detectable quantities of AIB. Ice samples dating from 1908 showed no traces of it, indicating that the Tunguska object did not deliver an appreciable organic signal to the earth.

In sediments from the KT boundary, we measured about 0.00005 gram of AIB for every square centimeter of the KT-boundary surface. If similar amounts of AIB were distributed over the entire surface of the earth (a generous assumption), then, in chemical terms, they would have created a two-billionths molar solution of AIB. That is like stirring a teaspoonful of sugar into a six-foot-deep swimming pool the size of a football field—too dilute a soup, in my opinion, for any kind of organic chemistry. The 4,500-year-old event recorded in the ice would have created a similarly dilute AIB solution. Even if (as Sagan and his colleagues have estimated) cosmic debris struck the prebiotic earth at 10,000 times the present levels, the resultant prebiotic soup would still have been much too weak, I believe, to engender life.

AT BEST, THEN, THE LIFE-FROM-SPACE HYPOTHesis comes off as a desperate attempt to explain how life could have formed in a fundamentally hostile environment. It would make far more sense to take the obvious given—that life exists—and ask under what conditions it was likely to form. As the Miller–Urey experiment showed, a reducing environment gives rise to organic molecules in concentrations orders of magnitude higher than does even the most optimistic alternative scenario. The logical follow-up is to ask another question: How could such an environment have come about?

The scenario we propose rests on a small but critical leap of logic. The Miller–Urey experiment assumed that the key ingredients for any earthly prebiotic soup must have originated in the earth's atmosphere. The most important reactions in its abiotic synthesis, however, took place not in air but in water. The amino acids in Miller's reacting vessels, for instance, came about via a reaction pathway known as the Strecker synthesis, developed in 1850 by the German–Norwegian chemist Adolph Friedrich Ludwig Strecker: gaseous hydrogen and methane reacted to produce slightly more complex molecules called aldehydes and ketones, and hydrogen cyanide, a lethal poison today. Those molecules then reacted with ammonia *in water* to produce amino acids. The reducing environment needed to pave the way for life was not the atmosphere but the ocean.

We propose that on the prebiotic earth, the atmosphere and the ocean were divided at least sporadically into two separate reacting vessels. How could that have come about? Easily. Suppose the early atmosphere had the "wrong" composition, too thin a blanket of greenhouse gases to keep the warmth of the faint young sun from leaking away into space. Suppose the ocean froze. Suppose, contrary to the conventional wisdom, that the whole earth was sheathed in ice. What then?

For one thing, the ocean would not have frozen all the way to the bottom. Even today, radioactive decay in the earth's interior causes heat to flow outward through the ocean-bottom, warming the waters at a rate of about eight-hundredths of a watt for every square meter of ocean floor. Four billion years ago the heat flow may have been about three times as strong. If the surface temperature of the ice was minus forty degrees Celsius, a simple calculation balancing the heat flow from the ocean floor to the atmosphere shows that the ice layer would have been about 300 meters thick. Most of the ocean, to an average depth of several kilometers, would have stayed liquid (albeit at a bracing minus two degrees Celsius).

The icy water could easily have steeped into prebiotic broth. Methane and ammonia bubbling up through hydrothermal vents could have provided the raw ingredients for processes such as the Strecker synthesis. The cold would have prolonged the life spans of organic molecules, making it much more likely that a chemical complex would participate in a productive reaction. The important organic building block hydrogen cyanide, for instance, lasts an average of 1,000 years at fifty degree Celsius but a hundred times as long at zero degrees Celsius. Trapped under the ice, the molecules would accumulate year after year, century after century, forming a rich fertilizer for organic synthesis.

Under such a scenario, what goes on above the ice becomes much less important than what goes on underneath it. Theorists may pick and choose from a wide variety of possible atmospheres, ranging from a classical Miller–Urey mixture to the currently favored carbon dioxide–nitrogen, only with much more moderate levels of carbon dioxide. Thinking the unthinkable has its rewards.

WHAT ABOUT THE ALBEDO EFFECT? A FEW paragraphs ago I said that a frozen ocean would stay frozen forever. Yet only last week I left my office in La Jolla and went running on the beach, and the ocean was anything but frozen. What drove away the ice?

Only one answer is possible, and it arrives trailing a pleasing dose of irony. Cosmic collisions could have done the job. According to a model by the geophysicist Norman Sleep of Stanford University and his associates, when an extraterrestrial object strikes the earth, 75 percent of its kinetic energy ends up buried at the site of impact or dispersed into space: the rest of the energy heats the atmosphere. A simple calculation shows that a chunk of debris a hundred kilometers across could have melted all the ice on earth. It is possible that a much smaller body—similar to the one that struck the earth at the KT boundary—could have melted the ice, if it made a large enough hole. Thanks to the albedo effect, such a hole would gather heat, widen, suck in more heat and widen further, triggering a runaway melt.

If a global ice shell did exist, there is no way of knowing the exact circumstances of its demise. Perhaps one impact was enough to melt it forever. Perhaps the earth passed through several cycles of impacts and freezes, each new melt enabling a fresh burst of gases to bubble out of solution into the atmosphere: each new freeze ratcheting the prebiotic soup toward greater chemical sophistication. In that case, the vision of life under ice starts to look a little like panspermia-through-the-looking-glass, with extraterrestrial visitors cast not as instigators but as liberators. It also makes a nice coda to the impact-frustration hypothesis, framing the emergence of the first organisms in the silence between two cosmic drum rolls: a long one that boils the oceans and delays the advent of life: then a short one that thaws the ice and sets life free. •

JEFFREY L. BADA is a professor of marine chemistry at the Scripps Institution of Oceanography, University of California, San Diego.

Article 4

Nature's feet of clay

Many creation myths involve humans being made from clay such as Adam in *Genesis*, or the creation stories of the American Indians. Now it seems that clay did play a big part in the origin of life on earth.

Scientists' belief in an omnipotent Creator — clay modeller or not — took a back seat after the publication of *Origin of species*. Darwin speculated that life began in 'a warm little pond', where all the necessary chemical components could come together — the 'primordial soup' model. However, there is an unsolved problem about how chemical evolution managed the leap from simple organic chemicals such as amino acids to the complex biochemistry of living organisms.

The main problem with the simple prebiotic soup model is that hydrolyis limits the length of polymers that can self condense in solution. This prevents RNA and DNA from forming unaided. Some kind of stepping stone or scaffolding is needed. J. D. Bernal suggested that clay minerals might provide the necessary template, an idea enlarged upon by Graham Cairns Smith (*Genetic takeover*, Cambridge: CUP, 1982; *Chem. Br.*, October 1979, p576). Last month a group including Leslie Orgel (Salk Institute, US), a veteran of many years' study on chemical evolution, published research showing that mineral surfaces can indeed encourage longer chains to form (J. P. Ferris *et al.*, *Nature*, 1996, **381**, 59).

The experiments they reported are reminiscent of the solid phase techniques used in DNA sequencing and combinatorial chemistry. Whereas only short oligomers (typically no more than a 10 mer) will be produced in solution, Orgel and his colleagues constructed 55 mers by repeatedly 'feeding' the mineral with monomers. Montmorillonite clay is an effective catalyst for nucleotide oligomerisation, while illite clay and hydroxyapatite do the same for amino acids.

Orgel and his coworkers claim that 'the "feeding" protocol that we have outlined is a plausible model of prebiotic polymerization. The repeated incubations with low concentrations of activated monomers simulate conditions in environments where rocks are in constant contact with low concentrations of activated substrates or are episodically washed with higher concentrations'.

They suggest that subject to the vagaries of hydrolysis their model could lead to mineral surfaces being covered by polymers of 'enormous length'. This would result in a slimy coating on the rocks — just the environment where some of the earliest fossils in Earth history have been found.

Extract 1

Search for Past Life on Mars: Possible Relic Biogenic Activity in Martian Meteorite ALH84001

David S. McKay, Everett K. Gibson Jr.,
Kathie L. Thomas-Keprta, Hojatollah Vali,
Christopher S. Romanek, Simon J. Clemett,
Xavier D. F. Chillier, Claude R. Maechling, Richard N. Zare

Fresh fracture surfaces of the martian meteorite ALH84001 contain abundant polycyclic aromatic hydrocarbons (PAHs). These fresh fracture surfaces also display carbonate globules. Contamination studies suggest that the PAHs are indigenous to the meteorite. High-resolution scanning and transmission electron microscopy study of surface textures and internal structures of selected carbonate globules show that the globules contain fine-grained, secondary phases of single-domain magnetite and Fe-sulfides. The carbonate globules are similar in texture and size to some terrestrial bacterially induced carbonate precipitates. Although inorganic formation is possible, formation of the globules by biogenic processes could explain many of the observed features, including the PAHs. The PAHs, the carbonate globules, and their associated secondary mineral phases and textures could thus be fossil remains of a past martian biota.

Extract 2

In examining the martian meteorite ALH84001 we have found that the following evidence is compatible with the existence of past life on Mars: (i) an igneous Mars rock (of unknown geologic context) that was penetrated by a fluid along fracures and pore spaces, which then became the sites of secondary mineral formation and possible biogenic activity; (ii) a formation age for the carbonate globules younger than the age of the igneous rock; (iii) SEM and TEM images of carbonate globules and features resembling terrestrial microorganisms, terrestrial biogenic carbonate structures, or microfossils; (iv) magnetite and iron sulfide particles that could have resulted from oxidation and reduction reactions known to be important in terrestrial microbial systems; and (v) the presence of PAHs associated with surfaces rich in carbonate globules. None of these observations is in itself conclusive for the existence of past life. Although there are alternative explanations for each of these phenomena taken individually, when they are considered collectively, particularly in view of their spatial association, we conclude that they are evidence for primitive life on early Mars.

Europa: Prospects for an Ocean and Exobiological Implications

J. Oro, S. W. Squyres,
R. T. Reynolds and T. M. Mills

Extract 1 (Activity 3.5)

Introduction

As far as we know, Earth is the only planet in our solar system that supports life. It is natural, therefore, that our understanding of life as a planetary phenomenon is based upon Earth-like planets. For example, Mars has received a great deal of attention as a possible former abode for life because its climate appears to have been somewhat more Earth-like in the past. However, there are environments in the solar system where liquid water, commonly believed to be a prerequisite for biological activity, may exist in a distinctly non-Earth-like environment. One such location is Europa, one of the Galilean satellites of Jupiter. The possibility that liquid water exists on Europa presents us with some interesting exobiological implications concerning the potential of the satellite to support life.

An Ocean on Europa?

The images of Europa taken during the Voyager flybys of the Jovian system show a very bright surface transected by a network of long linear features of lower albedo and brownish colour. The features are seen down to the resolution of the images (~4 km/line pair) and are as large as tens of kilometers across with maximum lengths comparable to the radius of Europa (~1500 km). Their geometry indicates that they are probably fractures caused by extensional stresses in Europa's crust. Other linear topographical features include low ridges of unknown origin. Overall, however, the surface of Europa is remarkably level. Only a handful of small impact craters have been positively identified, and it is doubtful that topography anywhere on the satellite exceeds a few hundred meters.

Even the most basic pre-Voyager information about Europa appeared enigmatic. The surface is composed primarily of H_2O and frost, yet the satellite's density clearly indicates that it is predominantly a silicate body. Simple but powerful cosmochemical arguments suggest that all of the Galilean satellites should be composed primarily of silicates and varying amounts of H_2O. Water frost was first positively identified as a major component of Europa's surface in 1972, based on the presence of strong infrared H_2O absorption features in its reflectance spectrum. This finding has been confirmed by later observations, which have led to the conclusion that more than 95% of the spectroscopically detectable material on the surface of Europa is water. Despite the nearly pure water ice surface composition, the density of Europa is known to be 2.97 g cm^{-3}. A density this large indicates a composition dominated by silicates, with only a relatively small admixture of H_2O. If the silicate component of Europa is largely dehydrated and has a density like that of Io (3.57 g cm^{-3}), then Europa is composed of 7 to 8% free H_2O by mass. If, however, the silicates are hydrated, then the density is consistent with a composition including little or no free H_2O.

Three models have been proposed for the the internal structure of Europa that are consistent with the satellite's surface composition and density. These are the *thin ice, thick ice,* and *ice/ocean* models (fig. 6-1). The thin ice model suggests that the silicates in Europa's interior are largely hydrated and that the surface H_2O ice is a very thin layer (a few km) lying over the hydrated silicates. In the thick ice model, enough internal heat is produced and retained to dehydrate the silicates, driving the H_2O to the surface to form a layer of solid ice on the order of 100 km thick. In the ice/ocean model still more heat is produced and retained, melting much of the ice. The internal structure of Europa in this model consists of a dehydrated silicate interior, an ocean of liquid water on the order of 100 km thick, and a thin (10 km) ice layer on the surface.

All of these models are consistent with Europa's density and surface composition and choices among them must be made on the basis of surface morphology and models of Europa's internal thermal evolution.

End of Extract 1

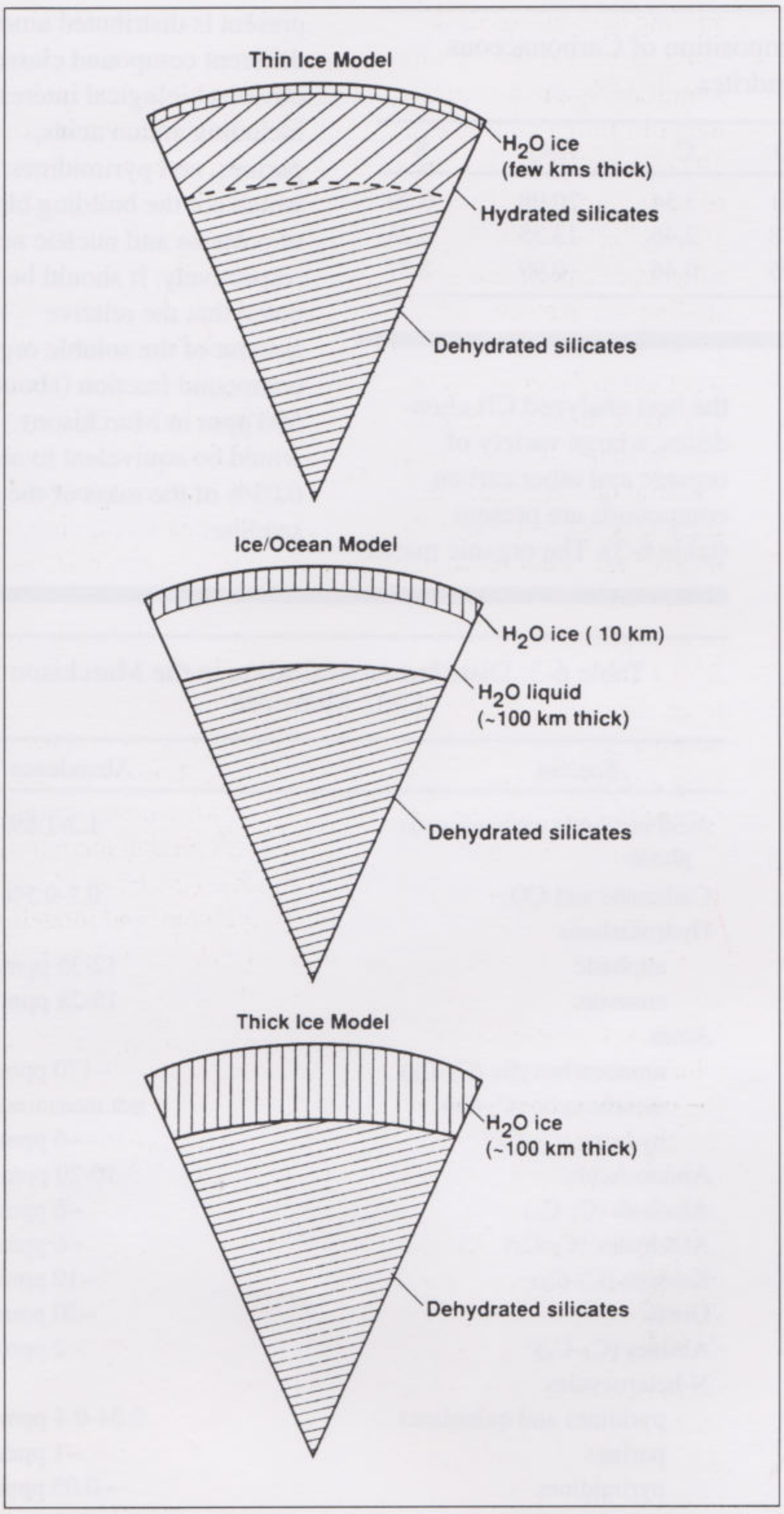


Figure 6-1. Three models proposed for the internal structure of Europa include the thin ice, thick ice, and ice/ocean models.

Extract 2 (Activity 3.6)

Table 6-2: Chemical Composition of Carbonaceous Chondrites

	SiO_2	MgO	C	H_2O	S
Type I	22.56	15.21	3.54	20.08	6.20
Type II	27.57	19.18	2.46	13.35	3.25
Type III	33.58	23.75	0.46	0.99	2.27

If we limit our discussion to the case of Europa and compare it with Type II carbonaceous chondrites (CII), we see that in addition to sulfur (3.25%) the CII chondrites contain relatively large amounts of water (13.35%) and carbon compounds (2.46%) (table 6-2). This shows that planetesimals of this composition, which are believed to have been involved in the formation of Europa through collisions with the proto-satellite throughout its early history, would have carried with them more than sufficient amounts of H_2O to have provided the calculated water fraction of Europa which is at least 7% of the mass of the satellite ($M_e = 4.87 \times 10^{25}$ g). But what is more important from an exobiological point of view is that a significant presence of organic compounds in Europa is also suggested by this comparison. What are the natures of the organic compounds? As indicated by the studies done on the Murchison meteorite, one of the best analyzed CII chondrites, a large variety of organic and other carbon compounds are present (table 6-3). The organic matter present is distributed among different compound classes of obvious biological interest, including amino acids, purines, and pyrimidines, which are the building blocks of proteins and nucleic acids, respectively. It should be noted that the relative amount of the soluble organic compound fraction (about 500 ppm in Murchison) would be equivalent to about 0.05% of the mass of the satellite.

Table 6-3: Distribution of Carbon in the Murchison CM2 Meteorite

Species		Abundance
Acid insoluble carbonaceous phase		1.3-1.8%
Carbonate and CO_2		0.1-0.5%
Hydrocarbons		
aliphatic		12-35 ppm
aromatic		15-28 ppm
Acids		
monocarboxylic (C_2-C_8)		~170 ppm
dicarboxylic (C_2-C_9)		not measured
hydroxy (C_2-C_5)		~6 ppm
Amino Acids		10-20 ppm
Alcohols (C_1-C_4)		~6 ppm
Aldehydes (C_2-C_4)		~6 ppm
Ketones (C_3-C_5)		~10 ppm
Ureas		~20 ppm
Amines (C_1-C_4)		~2 ppm
N-heterocycles		
pyridines and quinolines		0.04-0.4 ppm
purines		~1 ppm
pyrimidines		~0.05 ppm
poly-pyrroles		<<1 ppm
	Sum:	1.43-2.35%
	Total carbon:	2.0-2.5%

We can translate the CII meteorite data into a possible composition of the upper hydrous layers of Europa. Assuming that about 90% of the satellite's water is in the liquid phase (namely 6.3% of the mass of the satellite) the concentration of organic compounds dissolved in such a subsurface ocean would be of the order of 1%. In absolute terms the amounts of total carbon (carbonates plus carbon compounds) in Europa would be 1.2×10^{24} g and that of soluble organic compounds 2.3×10^{22} g. The latter figure is at least four orders of magnitude higher than the total organics in Earth's active biosphere. Moreover, a concentration of 1% organics in water is a sufficient concentration to lead to significant reaction rates for abiotic biochemical synthesis under favourable conditions. Whether these reactions could have led to emergence and evolution of life on Europa is not known and would have depended very much on the conditions of the satellite toward the end of the accretion phase, as discussed earlier.

Considering what may be occurring today on the surface of Europa, possible evidence for the synthesis of organics may be visible on the satellite's surface. The presence of brown colorations on the surface fractures of Europa has led to the speculative suggestion that the colorations present are due to organics produced abiotically on or below the surface of Europa. Organics synthesized below the surface could be brought to the surface as the result of fractures in the ice. However, at this time the results from a number of investigations that in some way model the conditions on Europa have produced only highly speculative implications.

End of Extract 2

Energy Sources for Biosynthesis and Metabolic Activity

Assuming Europa's ocean exists, there are primarily two major energy sources in the aqueous layers of Europa which could be utilized biologically for the synthesis of biochemical compounds and for different metabolic reactions.

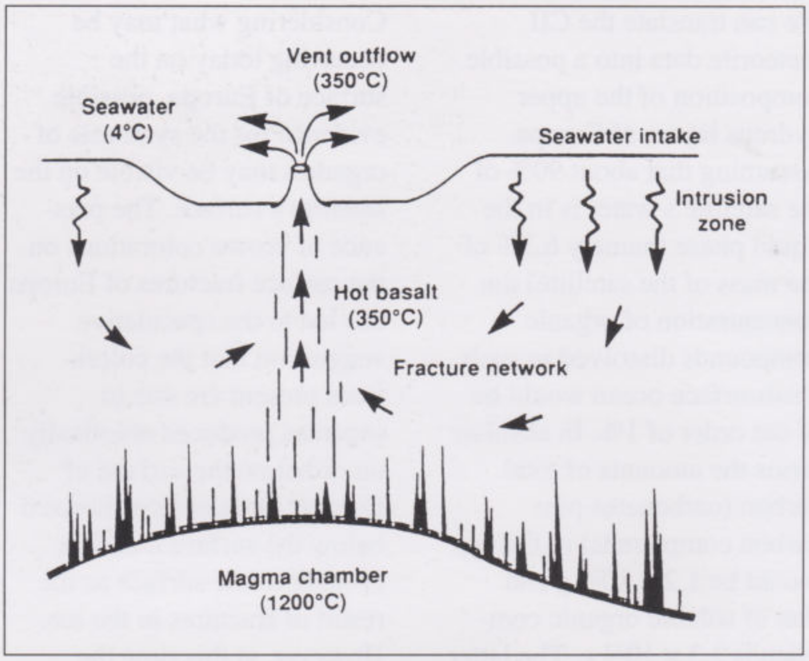


Figure 6-3. On Earth, the concentration of thermal energy in local high-temperature regions at submarine crests has led to large-scale hydrothermal activity. Reactions between sea water and hot basalt beneath the ocean floor produce solutions that ascend to the seafloor and merge as hot springs whose sulfur and hydrogen compounds sustain chemosynthetic or chemolithotrophic bacteria.

The first is the dissipation of tidal energy in the form of heat, which presumably maintains a significant fraction of the water on Europa as liquid. In Io, the dissipation of tidal energy does not occur in a very uniform manner, but rather in discrete zones or spots, as shown by the more than 100 sulfur volcanoes on its surface. At a much lower scale the transfer of heat from the core of Europa to the bottom liquid water layers of its ocean may also occur in localized places. The existence of these non-uniform areas of heat transfer in Europa are at present only speculative, however, assuming they exist, they could be analogs of terrestrial deep oceanic thermal vents. On the Earth, the concentration of thermal energy in local high-temperature regions at submarine crests has led to large-scale hydrothermal activity (fig. 6-3). Reactions between sea water and hot basalt beneath the ocean floor take place at temperatures in excess of 300°C. The resulting solutions ascend to the sea floor where they emerge as hot springs that sustain oases of life. At the base of the food chain in these oases are chemosynthetic or chemolithotrophic bacteria that derive their entire energy supply from the reaction of geothermally produced sulfur and hydrogen compounds released from the vents.

Analogous hydrothermal regions conceivably could exist on the bottom of Europa's ocean. The calculated mean heat flux from Europa's silicate interior is less than a third of the mean heat flow of the Earth. In fact, it is only some 50% higher than that of the Moon, which is not observed to have any volcanic activity at the present. However, since the Moon's near-surface regions are virtually water-free this does not rule out the possibility of hydrothermal activity on Europa. Current data and modeling techniques are insufficient to provide a quantitative assessment of the probability of such activity.

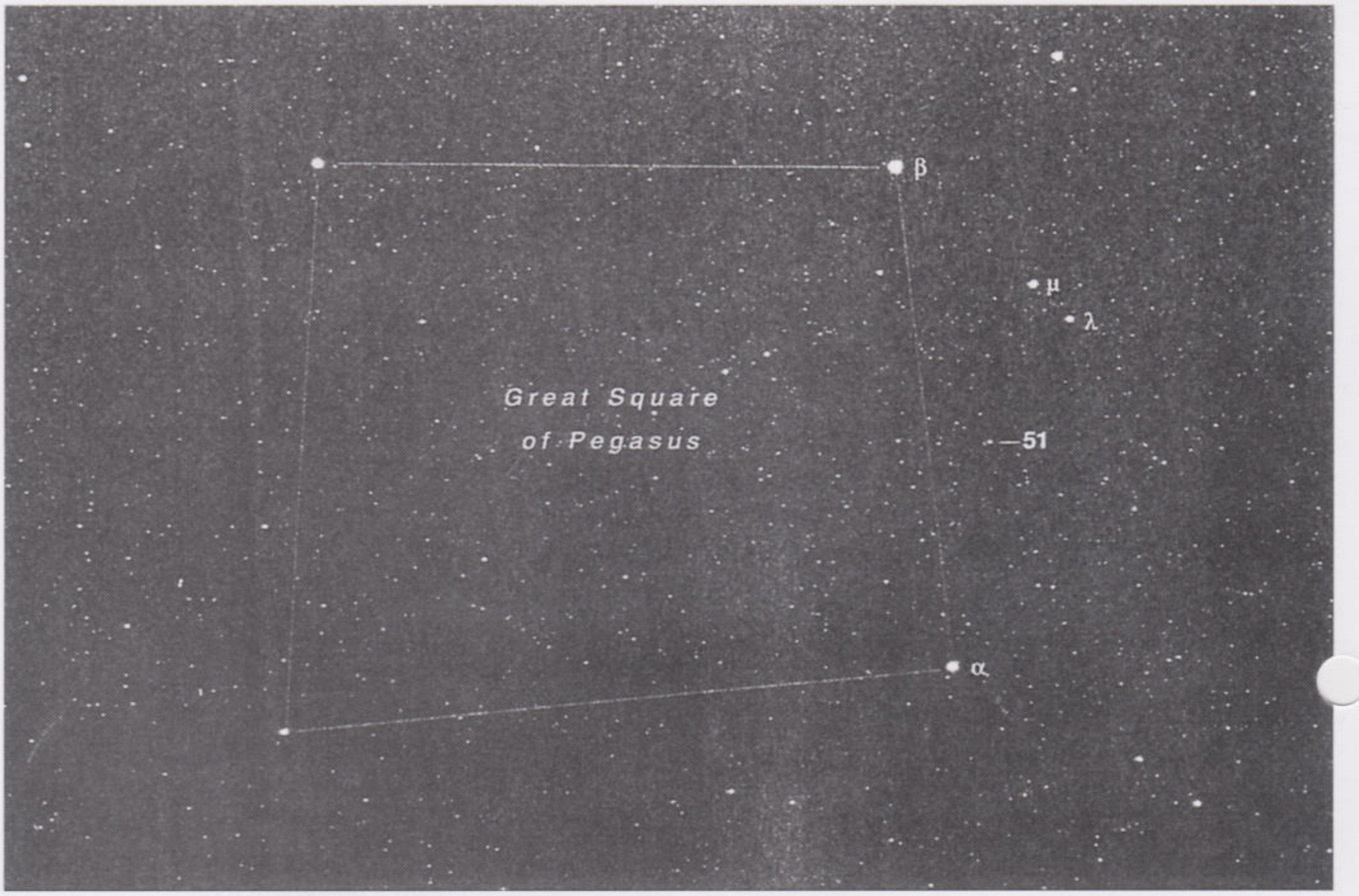
The existence of sulfur on Europa, as suggested by the carbonaceous chondrite model, offers an additional similarity between terrestrial and theoretical Europan oceanic vents. It is of interest to note in this respect that due to the unequilibrated nature of the matter in these meteorites, and also probably in the Jovian nebula, several of the different states of oxidation of the sulfur compounds may coexist simultaneously, e.g., from the most reduced sulfides to the most oxidized sulfates. This would establish an adequate chemical redox potential ideally suited for utilization in metabolic reactions early in the satellite's development. The latter infers the plausibility of the existence, at some time during Europa's past or present history, of a microbial ecosystem analogous to those that exist presently near terrestrial submarine geothermal vents.

The second form of energy that could be utilized biologically is the solar radiation which reaches the icy surface of the satellite. Assuming that solar energy would reach liquid water only where the ice crust had recently been fractured, it is straightforward to predict the maximum amount of light that could reach the Europan ocean. Calculations show that the light levels in the water would be extremely low, even immediately after a fracturing event. However, for up to 4.5 years after a fracturing event they could remain higher than the levels at which photosynthesis occurs in primitive microbial mats in perennially ice-covered lakes in the Dry Valleys of Antarctica. So, if such fracturing does indeed occur, very limited local environments may exist for extremely short periods of time that are within the range of adaptation of simple Earth organisms.

Two very important caveats should be added to this discussion of solar energy input. First, the scenario presented here represents a strict upper limit in that it does not consider loss of thermal energy to the walls of a fracture. If a fracture is narrow relative to the thickness of the forming ice cover, heat loss to the walls will cause freezing that may be substantially more rapid than we calculate. Second, we do not know the frequency with which fracturing events occur, or indeed if they take place at all. Using the most liberal estimates of fracturing, based on an inferred resurfacing rate by recondensing frost, we obtain a maximum total fracture area of approximately 5 km^2/yr, which is an exceedingly small fraction of the total surface area of the satellite.

End of Extract 3

Article 9



Plucked from obscurity to make astronomical history, 51 Pegasi shines demurely at magnitude 5.5 just west of the Great Square of Pegasus. Seen in binoculars or a small telescope, it displays the yellow-white color of the Sun — which is in fact almost its twin. North is up, and the field shown is about 35° wide. Photograph by Akira Fujii.

The first planet found orbiting another Sun-like star is not at all what astronomers expected.

The Planet of 51 Pegasi

By Alan M. MacRobert and Joshua Roth

A DISCOVERY that most astronomically minded people have awaited since childhood — one that many expected to go to their graves never seeing — has finally happened. Unless astronomers are badly mistaken, the first planet orbiting a normal star other than the Sun has at last been found. And what a strange planet it is turning out to be.

For centuries, one of astronomy's greatest unknowns has been whether other stars have solar systems. Generations of textbooks have taught why finding planets around even the nearest stars was impossible. Any planet would be much too dim and close to its star's overwhelming glare to be seen directly.

But in recent years, searches for massive planets — "Jupiters" and up — have become possible by indirect means. For instance, if a planet orbits a star, the star does not stay still. Both bodies revolve around their common center of gravity, or *barycenter*, with the sizes and speeds of their orbits proportional to the ratio of their masses. One promising planet-search strategy has been to look for these wobbling motions directly by monitoring the exact positions of stars (*S&T:* October 1987, page 360). Such astrometric approaches have yet to bear fruit.

A second approach is to look for the slight changes that the wobbling would cause in the star's radial velocity, its motion toward or away from Earth. For example, Jupiter has about a thousandth the mass of the Sun and an orbital speed of about 13,000 meters per second. Thus the Sun's corresponding movement, or *reflex motion*, is 13 meters per second — the speed of a leisurely drive through town.

About a half dozen groups of astronomers around the world have built highly specialized spectrographs that can measure stellar radial velocities as small as a few dozen or even just a few meters per second. The instruments look for extremely tiny Doppler-induced wavelength shifts in a star's spectral lines. This strategy favors massive planets and those in tight orbits around their stars, since these will induce the fastest reflex motions.

One ambitious radial velocity program is being carried out by Michel Mayor and graduate student Didier Queloz (Geneva Observatory). Since April 1994 they

have been tracking the radial velocities of 142 selected Sun-like stars that they had previously found to be single and hence likelier to harbor planets. Using a state-of-the-art spectrograph on the 1.9-meter telescope at Haute-Provence Observatory in France, they can achieve a velocity resolution of 12 meters per second. This was good enough, they reasoned, to identify stars having planets several times the mass of Jupiter.

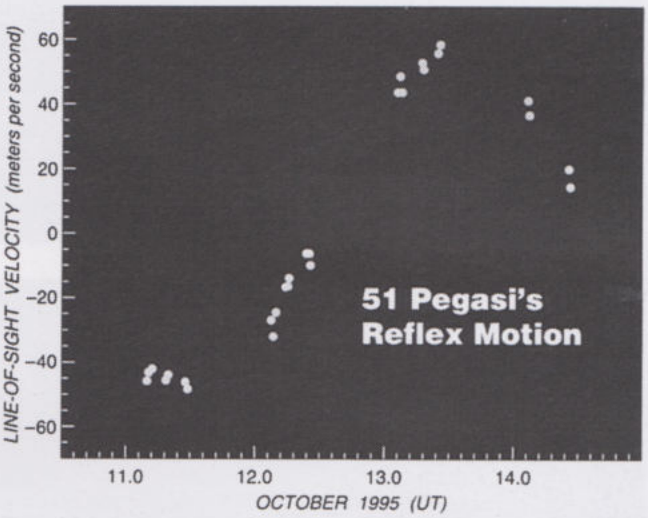
But when Mayor and Queloz hit the jackpot, it was not in a way anyone expected. In early October they announced that a planet with at least half the mass of Jupiter is orbiting 51 Pegasi, a type-*G*2–3 main-sequence star. The orbital period was an amazingly short 4.229 ± 0.001 days. The star's radial velocity was oscillating with a semiamplitude of about 60 meters per second (120 meters per second peak to peak), and its period had remained rock-steady for more than a ar.

Shining at magnitude 5.5 from 40 light-years away, 51 Pegasi is nearly a dead ringer for the Sun. Like the Sun it is a slow rotator at least a few billion years old. The fast orbital period means the planet must circle the star at a distance of only 7 million kilometers — eight times closer than Mercury is to the Sun, but still very far from the star's surface. The planet should be heated to more than 1,000° Celsius by a stellar disk appearing 10° wide in its sky — the width of a tennis ball held at arm's length.

COOL STARS, HOT NEWS

Mayor broke the news on October 6th during a workshop on cool stars in Florence, Italy. Within days other astronomers using even higher-resolution equipment were rushing to check on 51 gasi for themselves. Among them ere Geoffrey Marcy (San Francisco State University) and Paul Butler (University of California, Berkeley), who run what is probably the most precise stellar radial-velocity program in the world. Their spectrograph at Lick Observatory's 3-meter Shane reflector can measure stellar Doppler shifts to an accuracy of 3 meters per second.

Ironically, Marcy and Butler had excluded 51 Pegasi from their own ongoing search program because the Yale *Bright Star Catalogue* misidentifies it as a subgiant. On hearing of the Swiss team's discovery, they lost no time using four nights of previously scheduled telescope time — nearly a full "year" for the new planet — to obtain 27 high-precision measurements. These showed the star indeed wobbling with a semiamplitude of 51 ± 2 meters per second.



The line-of-sight motions that betray 51 Pegasi's planet are dramatically clear in this data set from four nights of spectroscopy at Lick Observatory's 3-meter reflector on Mount Hamilton, California. Several teams have now confirmed that the star approaches Earth by about 51 meters per second faster, then slower, than average. The data fit a perfect sine curve, indicating a circular orbit. Courtesy Geoffrey Marcy.

Moreover, the data points fit a perfect sine curve, implying that the planet's orbit is circular — as astronomers knew it would have to be if the planet were real. Tidal effects should circularize the orbit of a massive planet that's so close to its star; the energy loss that occurs when tides are raised tends to make eccentric orbits round.

By press time at least two other groups of astronomers had also confirmed the star's sinusoidal velocity curve.

ALTERNATIVE EXPLANATIONS

But are these minuscule velocity variations really the mark of an orbiting planet? Might something else be causing them?

"At the start," recalls Mayor, "we were not at all excited" by the radial-velocity variations, because he and Queloz presumed that something other than a planet must be at work. The last two decades have seen many premature planet "discoveries" withdrawn or shot down, starting with claims from the 1940s to 1970s of wobbling positions for several nearby red dwarfs. The only extrasolar planets that have withstood scrutiny are the three recently found to orbit PSR 1257+12, a pulsar 1,600 light-years away in Virgo (*S&T:* May 1992, page 493, and May 1994, page 10).

But as the days went by and astronomers raced to gather (or release) new information, the case for 51 Pegasi's planet grew stronger. No pulsations in a solar-type star, for instance, should have anywhere near a 4-day period. The Sun's fundamental oscillation period is 5 minutes, and even the longest possible periods that solar seismologists are searching for in the Sun are just a few hours long.

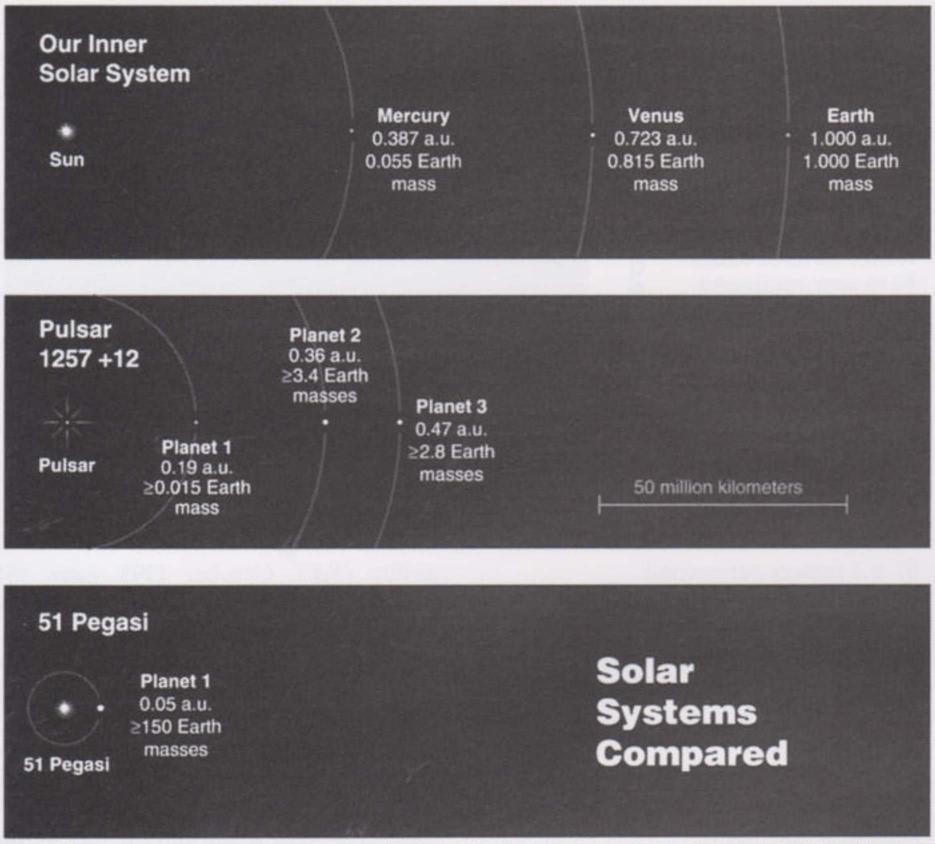
The period of 51 Pegasi has remained as constant as clockwork for more than a year, unlike many kinds of intrinsic variability (*S&T:* October 1995, page 15). Moreover, neither stellar pulsations nor other effects, such as modulation of spectral-line profiles by magnetic or starspot activity, would be expected to show such perfectly sinusoidal behavior.

Finally, in late October Geneva Observatory astronomers announced that intensive monitoring of 51 Pegasi's brightness had been carried out at the European Southern Observatory (ESO) — and that any variations in the star's brightness with a 4.2-day period larger than 0.002 magnitude could be ruled out. Regardless of any other effects, expansion and shrinking of the star's surface by 50 meters per second over a four-day cycle would result in brightness changes nearly 10 times that amount.

The data also enabled the astronomers to state that there was no evidence for eclipses in the system. Since a Jupiter-size planet would cause a 0.01 magnitude drop on crossing the face of 51 Pegasi, its orbital plane is presumably inclined to our line of sight by at least 5°.

The angle at which we view the orbit of the 51 Pegasi system is a critical unknown, for it sets the planet's mass. The widely quoted value of 0.5 Jupiter applies only if the orbit is seen within roughly 20° of edge-on. If the orbit is instead oriented nearly perpendicular to our line of sight, the radial velocity we see from Earth is only a small component of the star's actual reflex motion. The planet's mass could thus be a good deal larger. If we see the system almost pole-on — a slim possibility — the orbiting object could even be a brown dwarf, a failed star dozens of times Jupiter's mass.

Luckily some clues are at hand. The spectral lines of 51 Pegasi are broadened by Doppler shifts of 2 kilometers per second due to the star's rotation around its own axis. Again, this is only some unknown fraction of the star's true rotation



The three planetary systems currently known. The bottom two were found by slight radial velocity changes in their stars caused by the orbiting planets' gravitational tugs. The orbits are drawn to scale; at this scale the dots for the yellow stars would be somewhat smaller, and those for the planets and the pulsar would be microscopic.

speed, since we may not be facing its equator. But the lack of bright calcium lines in 51 Pegasi's spectrum implies that it has little surface activity — even a bit less than the Sun — and this is the sign of a star that rotates slowly (*S&T:* May 1995, page 12). Thus 51 Pegasi is unlikely to turn much faster than the 2 km per second actually seen, implying that we see the star not far from equator on. (Coincidentally, the Sun rotates at a similar rate.) Theory suggests that planets form near the plane of their star's equator, as happened in our own solar system. So it's a fair bet that the new planet's minimum mass of 0.5 Jupiter is close to the true value.

AN UNLIKELY WORLD

However, the very existence of a giant planet a mere 7 million km from its star throws current theories of planet formation into turmoil. Theorists have assumed that planets start coalescing when small, solid objects glom onto one another in a protoplanetary disk. If they become massive enough they draw in gas as well, becoming giants like Jupiter. However, Peter Bodenheimer (University of California, Santa Cruz), doubts that a solid core could ever coalesce so close to a young, hot, solar-mass star settling down to life on the main sequence.

Did the newfound planet instead form the way double stars do, as a separate condensation in the gas cloud that collapsed to form 51 Pegasi? Could it have formed farther out and somehow spiraled in, or was it flung in by a collision with another massive body? Did it start out much more massive, then erode or evaporate? These are among the questions theorists suddenly find themselves having to consider. In a scenario mentioned by Robert Noyes (Harvard-Smithsonian Center for Astrophysics), the "planet" might have started out with a star's worth of matter and dumped most of it onto 51 Pegasi early in their evolution. However, Noyes points out, the utter normalcy of 51 Pegasi argues against any such scheme.

The nature of the planet today is an even more fascinating mystery. It need not be the "giant iron wrecking ball" that Marcy first posited upon Mayor's announcement — an enormous, airless Mercury with five times the diameter of Earth. On the contrary, says George Wetherill (Carnegie Institute), it could retain a massive, Jupiter-style atmosphere despite 51 Pegasi's blazing proximity. But how it could have kept an atmosphere during the early T Tauri phase of the star's life, when 51 Pegasi was probably 100 times more luminous and blowing off a massive stellar wind, does give theorists pause.

A DISCOVERY TO SHARE

As this issue goes to press the floodgates are bursting with news of other possible planets and more massive substellar objects — brown dwarfs — in orbit around additional stars (see page 10) and possibly 51 Pegasi too. Astronomers engaged in patient, long-term searches are suddenly under pressure to release findings that are exciting but still partial and inconclusive. *Sky & Telescope* will continue to report on this unfolding story in the months to come. In the meantime, with Pegasus high in the northwestern sky each clear evening, both amateurs and professionals have a rare opportunity to share a momentous discovery with their neighbors and friends. As Marcy puts it, "Planets are something even a six-year-old can understand and get excited about."

Cautionary Tales

AMID THE EXCITEMENT surrounding the potentially momentous discovery of 51 Pegasi's planet, it's worth recalling the cases of extrasolar planets that were intially declared with fanfare, then debunked.

Barnard's Star, 61 Cygni, Lalande 21185, and several other red dwarfs that are among the closest stars to our solar system were once believed to show planet-induced wobbles. However, later astrometric attempts to confirm these findings disproved them instead (*S&T:* October 1987, page 360, and June 1995, page 68). VB8B, a substellar companion to Van Biesbroeck 8 allegedly seen in infrared images, was ruled out within two years of its much-heralded "discovery" (*S&T:* August 1987, page 139). And the first planet thought to orbit a pulsar resulted from a mistaken correction for the Earth's own motion around the Sun (*S&T:* May 1992, page 493).

On a positive note, at least as many viable candidates for other solar systems are currently under scrutiny and may add to what has become a new branch in astronomy's ever-growing family tree.

Typeset and printed in the United Kingdom by
Thanet Press Ltd, Union Crescent, Margate, Kent CT9 1NU
S103s435857i1.1

S103 B12A2

The Open University

S103 discovering science

Articles for Block 12: Part 2

Contents

(For Articles 1, 2, 3, 4, 6, 8 and 9 see Articles for Block 12: Part 1)

1.5 DS/P

SUP 43586 2

Figure 21 from Article 1

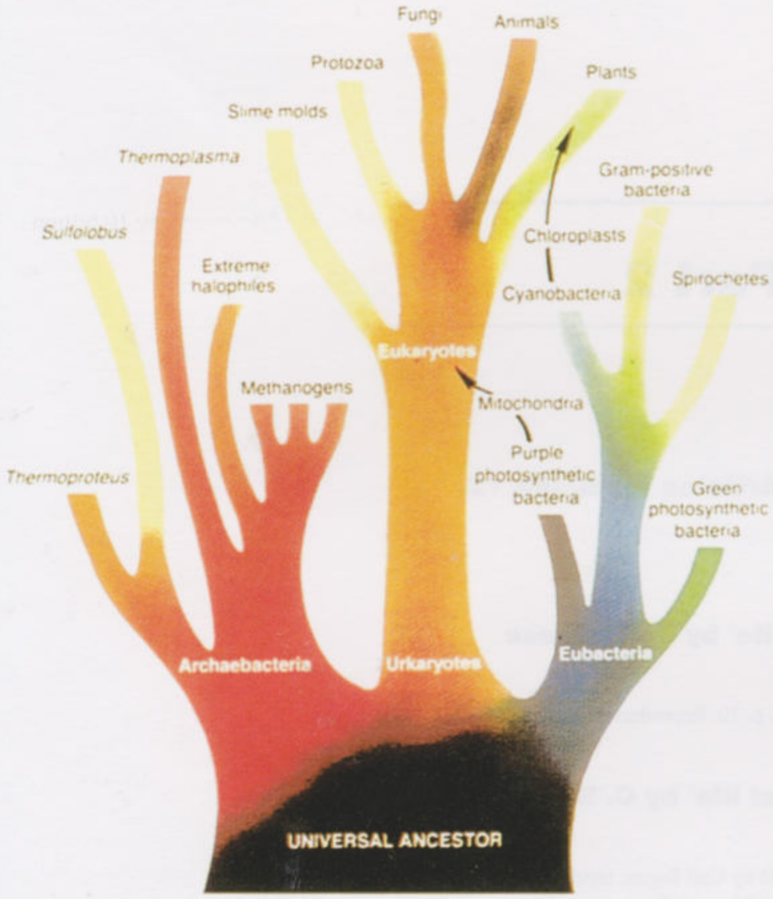


Figure 21. The universal phylogenetic tree. All known living forms fall into one of the three basic categories: the archaebacteria, the eubacteria, and the eukaryotes. The first two are prokaryotic cells; that is, they are small, do not have nuclear membranes or organelles, etc. However, the archaebacteria and eubacteria are not specifically related to one another. From fossil evidence bacteria appear to have been in existence for at least 3.5 billion years and perhaps longer, far before the advent of the eukaryotic groups. Each of the three major groups of organisms is quite distinct on the molecular level, differing from the others in the molecular details of their common cellular processes.

[For the text of Article 1, see *Articles for Block 12: Part 2*.]

Article 5

The Search for Extraterrestrial Life

The earth remains the only inhabited world known so far, but scientists are finding that the universe abounds with the chemistry of life

by Carl Sagan

Extract 1 (Activity 3.1)

In the past few decades the human species has begun, seriously and systematically, to look for evidence of life elsewhere. While no one has yet found living organisms beyond the earth, there are some reasons to be encouraged. Robotic space probes have identified worlds where life may once have gained a toehold, even if it does not flourish there today. The *Galileo* spacecraft found clear signs of life during its recent flight past the earth—a reassurance that we really do know how to sniff out at least certain kinds of life. And rapidly accumulating evidence strongly suggests that the universe abounds with planetary systems something like our own.

In practice, the community of scientists concerned with finding life elsewhere in the solar system has contented itself with a chemical approach. Human beings, as well as every other organism on the earth, are based on liquid water and organic molecules. (Organic molecules are carbon-containing compounds other than carbon dioxide and carbon monoxide.) A modest search strategy—looking for necessary if not sufficient criteria—might then begin by looking for liquid water and organic molecules. Of course, such a protocol might miss forms of life about which we are wholly ignorant, but that does not mean we could not detect them by other methods. If a silicon-based giraffe had walked by the Viking Mars landers, its portrait would have been taken.

Actually, focusing on organic matter and liquid water is not nearly so parochial and chauvinistic as it might seem. No other chemical element comes close to carbon in the variety and intricacy of the compounds it can form; liquid water provides a superb, stable medium in which organic molecules can dissolve and interact. What is more, organic molecules are surprisingly common in the universe. Astronomers find evidence for them everywhere, from interstellar gas and dust grains to meteorites to many worlds in the outer solar system.

Some other molecules—hydrogen fluoride, for example—might approach water in their ability to dissolve other molecules, but the cosmic abundance of fluorine is extremely low. Certain atoms, such as silicon, might be able to take on some of the roles of carbon in an alternative life chemistry, but the variety of information-bearing molecules they provide seems comparatively sparse. Furthermore, the silicon equivalent of carbon dioxide (silicon dioxide, the major component of ordinary glass) is, on all planetary surfaces, a solid, not a gas. That distinction would certainly complicate the development of a silicon-based metabolism.

On extremely cold worlds, where water is frozen solid, some other solvent—liquid ammonia, for instance—might be a key to a different form of biochemistry. At low temperatures, certain classes of molecules require very little activation energy to undergo chemical reactions, but because our laboratories are at room temperature and not, say, at the temperature of Neptune's satellite Triton, our knowledge of those molecules may well be inadequate. For the moment, though, carbon- and water-based life-forms are the only kinds we know or can even imagine.

On the earth the signature molecules of life are the nucleic acids (DNA and RNA), which constitute the hereditary instructions, and the proteins, which, as enzymes, catalytically control the chemistry of cell and organism. The codebook for translating nucleic acid information into protein structure is essentially identical for all life on the earth. This profound uniformity in the hereditary chemistry suggests that every organism on our planet has evolved from a common instance of the origin of life. If so, we have no way of knowing which aspects of terrestrial life are necessary (required of all living things anywhere) and which are merely contingent (the results of a particular sequence of happenstances that, had they gone otherwise, might have led to organisms having very different properties). We may speculate, but only by examining life elsewhere can biologists truly determine what else is possible.

The obvious place to start the search for life is in our own solar system. Robot spacecraft have explored more than 70 planets, satellites, comets and aster-

CARL SAGAN is noted as a researcher, lecturer and author. He received his Ph.D. in astronomy and astrophysics from the University of Chicago in 1960. In 1968 he joined Cornell University, where he is David Duncan Professor of Astronomy and Space Science and director of the Laboratory for Planetary Studies. He has participated in numerous National Aeronautics and Space Administration planetary missions, among them Viking, Voyager and Galileo. His research interests embrace the origin of life, the physics and chemistry of planetary atmospheres and surfaces, and the search for extraterrestrial intelligence. His books include *Cosmos; Contact*, a fictional account of a meeting between humans and an alien civilization; and the recently published *Pale Blue Dot: A Vision of the Human Future in Space.*

oids at distances varying from about 100 to about 100,000 kilometers. These ships have been equipped with magnetometers, charged-particle detectors, imaging systems, and photometric and spectrometric instruments that sense radiation ranging from ultraviolet to kilometer-wavelength radio. For the moon, Venus and Mars, observations from orbiters and landers have confirmed and expanded on findings transmitted back from flyby spacecraft.

None of these encounters has yielded compelling, or even strongly suggestive, indications of extraterrestrial life. Still, such life, if it exists, might be quite unlike the forms with which we are familiar, or it might be present only marginally. Or the remote-sensing techniques used for examining other worlds might be insensitive to the conceivably subtle signs of life on another world. The most elementary test of these techniques—the detection of life on the earth by an instrumented flyby spacecraft—had, until recently, never been attempted. The National Aeronautics and Space Administration's *Galileo* has rectified that omission.

Galileo is a dual-purpose spacecraft that incorporates a Jupiter orbiter and entry probe; it is currently in interplanetary space and is scheduled to reach the Jupiter system in December 1995. For technical reasons, NASA was unable to send *Galileo* on a direct course to Jupiter; instead the mission incorporated three gravitational assists—two from the earth and one from Venus—to send it on its journey. This looping course greatly lengthened the transit time, but it also permitted the spacecraft to make close-up observations of our planet. *Galileo*'s instruments were not designed for an earth-encounter mission, so circumstance fortuitously arranged a control experiment: a search for life on the earth using a typical modern planetary probe. The results of *Galileo*'s December 1990 encounter with the earth proved quite enlightening.

An observer looking at the data from *Galileo* would immediately notice some unusual facts about the earth. When my co-workers and I examined spectra taken by *Galileo* at near-infrared wavelengths (just slightly longer than red light), we noted a strong dip in brightness at 0.76 micron, a wavelength at which molecular oxygen absorbs radiation. The prominence of the absorption feature implies an enormous abundance of molecular oxygen in the earth's atmosphere, many orders of magnitude greater than is found on any other planet in the solar system.

Oxygen slowly combines with the rocks on the earth's surface, so the oxygen-rich atmosphere requires a replenishing mechanism. Some oxygen is freed when ultraviolet light from the sun splits apart molecules of water (H_2O), and the low-mass hydrogen atoms preferentially escape into space. But the great concentration of oxygen (20 percent) in the earth's dense atmosphere is very hard to explain by this process.

If visible light, rather than ultraviolet, could split water molecules, the abundance of oxygen could be understood, because the sun emits many more photons of visible light than of ultraviolet. But photons of visible light are too feeble to sever the H-OH bond in water. If there were a way to combine two visible light photons to break apart the water molecule, then everything would have a ready solution. Yet so far as we know, there is no way to accomplish this feat—except through life, specifically through photosynthesis in plants. The prevalence of molecular oxygen in the earth's atmosphere is our first clue that the planet bears life.

When *Galileo* photographed the earth, it found unmistakable evidence of a sharp absorption band painting the continents: some substance was soaking up radiation at wavelengths around 0.7 micron (the far red end of the visible spectrum). No known minerals show such a feature, and it is found nowhere else in the solar system. The mystery substance is in fact just the kind of light-absorbing pigment we would expect if visible photons were being added together to break down water and generate molecular oxygen. *Galileo* detected this pigment—which we know as chlorophyll—covering most of the land area of the earth. (Plants appear green precisely because chlorophyll reflects green light but traps the red and blue.) The prevalence of the chlorophyll red band offers a second reason to think that the earth is an inhabited planet.

Galileo's infrared spectrometer also detected a trace amount, about one part per million, of methane. Although that might seem insignificant, it is in startling disequilibrium with all that oxygen. In the earth's atmosphere, methane

What Is Life?

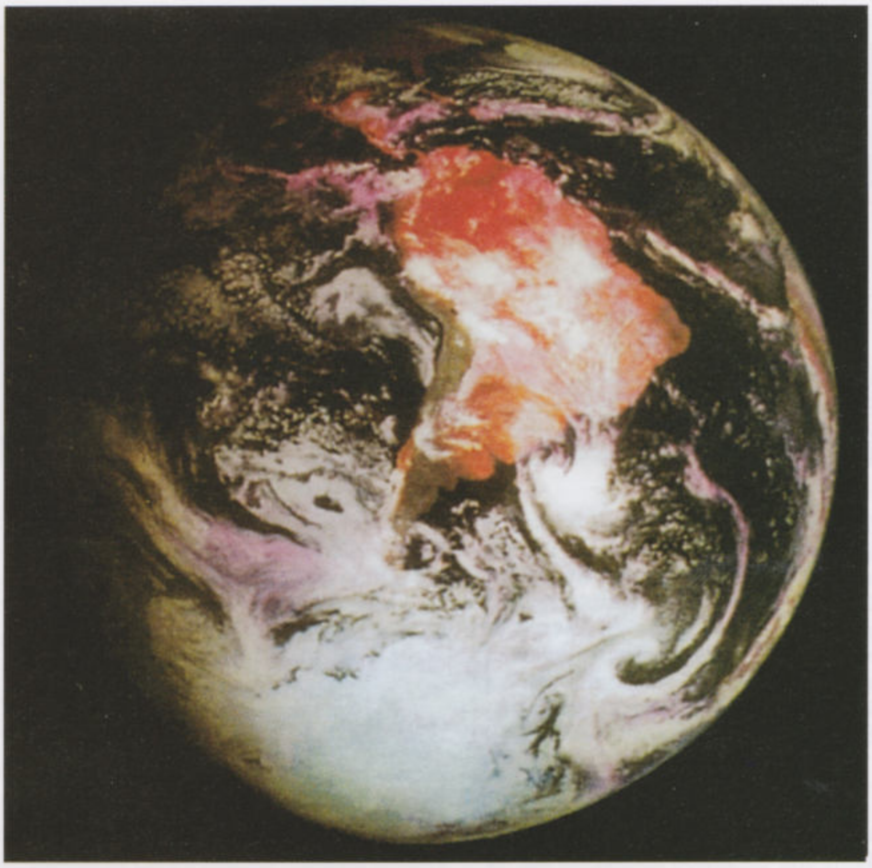
The search for extraterrestrial life must begin with the question of what we mean by life. "I'll know it when I see it" is an insufficient answer. Some functional definitions are inadequate: one might identify life as anything that ingests, metabolizes and excretes, but this description applies to my car or to a candle flame. Some more sophisticated definitions—for example, life as recognizable by its departure from thermodynamic equilibrium—fall afoul of the circumstance that much of nature (such as lightning and the ozone layer) is out of equilibrium.

Biochemical definitions—for example, defining life in terms of nucleic acids, proteins and other molecules—are clearly chauvinistic. Would we declare an organism that can do everything a bacterium can dead if it was made of very different molecules? The definition that I like best—life is any system capable of reproduction, mutation and reproduction of its mutations—is impractical to apply when we set down a spacecraft on another world: reproduction may not be done in public, and mutations might be comparatively infrequent.

rapidly oxidizes into water and carbon dioxide. At thermodynamic equilibrium, calculations indicate that not a single molecule of methane should remain. Some unusual processes (which we know to include bacterial metabolism in bogs, rumina and termites) must steadily refresh the methane supply. The profound methane disequilibrium is a third sign of life on the earth.

Finally, *Galileo*'s plasma-wave instrument picked up narrow-band, pulsed, amplitude-modulated radio emissions coming from the earth. These signals begin at the frequency at which radio transmissions on the earth's surface are first able to leak through the ionosphere; they look nothing like natural sources of radio waves, such as lightning and the earth's magnetosphere. Such unusual, orderly radio signals strongly suggest the presence of a technological civilization. This is a fourth sign of life and the only one that would not have been apparent to a similar spacecraft flying by the earth anytime within the past two billion years.

The Galileo mission served as a significant control experiment of the ability of remote-sensing spacecraft to detect life at various stages of evolutionary development on other worlds in the solar system. These positive results encourage us that we would be able to spot the telltale signature of life on other worlds. Given that we have found no such evidence, we tentatively conclude that widespread biological activity now exists, among all the bodies of the solar system, only on the earth.



LIFE ON THE EARTH betrays its presence in images and measurements made by the *Galileo* spacecraft. This false-color infrared image reveals a mysterious red-absorbing pigment (chlorophyll, which appears orange-brown here) painting the continents. No such pigment is seen anywhere else in the solar system. Spectra indicate that the earth's atmosphere is unusually rich in molecular oxygen and methane. *Galileo* has boosted scientists' confidence that they may be able to spot the telltale signs of life even if it is different from life on the earth.

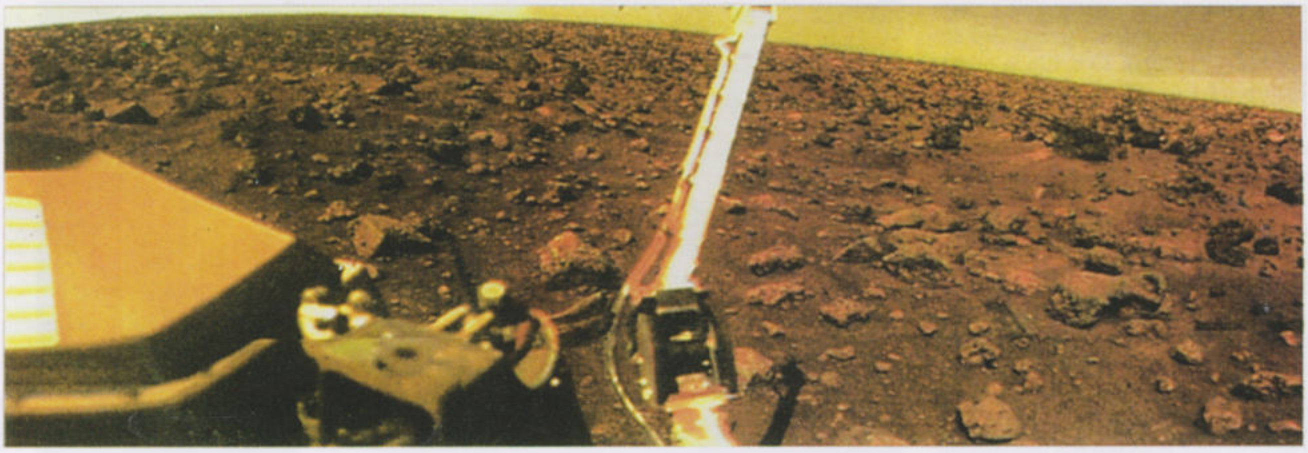
Mars is the nearest planet whose surface we can see. It has an atmosphere, polar ice caps, seasonal changes and a 24-hour day. To generations of scientists, writers and the public at large, Mars seemed the world most likely to sustain extraterrestrial life. But flybys past and orbiters around Mars have found no excess of molecular oxygen, no substances—whatever their nature—enigmatically and profoundly departing from thermodynamic equilibrium, no unexpected surface pigments and no modulated radio emissions. In 1976 NASA set down two *Viking* landers on Mars. I was an experimenter on that mission. The landers were equipped with instruments sensitive enough to detect life even in unpromising deserts and wastelands on the earth.

One experiment measured the gases exchanged between Martian surface samples and the local atmosphere in the presence of organic nutrients carried from the earth. A second experiment brought a wide variety of organic foodstuffs marked by a radioactive tracer, to see if there were life-forms in the Martian soil that ate the food and oxidized it, giving off radioactive carbon dioxide. A third experiment exposed the Martian soil to radioactive carbon dioxide and carbon monoxide to determine if any of it was taken up by microbes.

To the initial astonishment of, I think, all the scientists involved, each of the three *Viking* experiments gave what at first seemed to be positive results. Gases were exchanged; organic matter was oxidized; carbon dioxide was incorporated into the soil.

But there are reasons that these provocative results are not generally thought to provide a convincing argument for life on Mars. The putative metabolic processes of Martian microbes occurred under a wide range of conditions: wet and dry, light and dark, cold (only a little above freezing) and hot (almost the normal boiling point of water). Many microbiologists deem it unlikely that Martian microbes would be so capable under such varied conditions. Another strong reason for skepticism is that an additional experiment to look for organic molecules in the Martian soil gave uniformly negative results, even though the instruments could detect such molecules at a sensitivity of around one part per billion. We expected that any life on Mars—as with life on the earth—would be an expression of the chemistry of carbon-based molecules. To find no such molecules at all was daunting for the optimists among the exobiologists.

The apparent positive results of the life-detection experiments on the two *Viking* landers is now generally attributed to chemicals that oxidize the soil. These chemicals form when solar ultraviolet light irradiates the Martian atmosphere. A handful of *Viking* scientists still wonder whether extremely tough and resilient organisms might exist, so thinly spread over the Martian soil that their organic chemistry could not be detected but their metabolic processes could. Those scientists do not deny the presence of ultraviolet-generated oxidants, but they emphasize that nobody



***VIKING 2* LANDER scooped up bits of Martian soil and tested them for the presence of life and organic molecules. Despite tantalizing initial results, the *Viking* experiments suggest that Mars is, at present, a dead world. Future missions may search for fossils of organisms that might have lived billions of years ago, when Mars was warmer and wetter.**

has yet been able to explain fully the Viking life-detection results on the basis of oxidants alone. A few researchers have made tentative claims of finding organic matter in a class of meteorites (the SNC meteorites) that are thought to be bits of the Martian surface blasted into space during ancient impacts. More likely, the organic material consists of contaminants that entered the meteorite after its arrival on our world. So far there are no claims of discovering Martian microbes in these rocks from the sky.

For the moment, it is safe to say that *Viking* found no compelling case for life on Mars. No unambiguous signatures of life emerged from four very different, extremely sensitive experiments conducted at two sites 5,000 kilometers apart on a planet where fast winds transport fine particles around the globe. The *Viking* findings suggest that Mars is, today at least, a lifeless planet.

Could Mars have supported life in the distant past? The answer depends very much on how quickly life can arise, a topic about which we remain sadly ignorant. Astronomers are quite certain that, initially, the earth was inhospitable to life because of the collisions of planetesimals, the planetary building blocks that accreted together to form the earth. Early on, the earth was covered by a deep layer of molten rock. After that magma froze, the occasional arrival of large planetesimals would have boiled the oceans and sterilized the earth, if life had already arisen. Things did not calm down until about 4.0 billion years ago. And yet fossils reveal that by 3.6 billion years ago the earth abounded with microbial life (including large, basketball-size stromatolites, colonies of microorganisms). These early forms of life seem to have been biochemically very adept. Many were photosynthetic, slowly contributing to the earth's bizarre oxygen-rich atmosphere. Manfred Schidlowski of the Max Planck Institute for Chemistry in Mainz has studied carbon isotope ratios preserved in ancient rocks; that work provided (disputed) evidence that life was already flourishing 3.8 billion years ago.

The inferred time available for the origin of life on the earth is thus being squeezed from two directions. According to current knowledge, that amount of time may be as brief as 100 million years. When I first drew attention to this "squeeze"—in 1973, after lunar samples returned by *Apollo* clarified the chronology of impacts on the moon—I argued that the rapidity with which life arose on the earth may imply that it is a likely process. It is dangerous to extrapolate from a single example, but it would be a truly remarkable circumstance if life arose quickly here while on many other, similar worlds, given comparable time, it did not.

Between 4.0 and 3.8 billion years ago, conditions on Mars, too, may have favored the emergence of life. The surface of Mars is covered with evidence of ancient rivers, lakes and perhaps even oceans more than 100 meters deep. The Mars of 4.0 billion years ago was much warmer and wetter than it is today. Taken together, these pieces of information suggest, although they hardly prove, that life may have arisen on ancient Mars as it did on the ancient earth. If so, as Mars evolved from congenial to desolate, life would have held on in the last remaining refugia—perhaps saline lakes or places where the interior heat had melted the permafrost. Most planetary scientists agree that searching for chemical or morphological fossils of ancient life should have high priority in future Martian exploration. Although it is a long shot, searching for life in contemporary Martian oases might also be a productive endeavor.

End of Extract 2 (Activity 3.2)

It is now clear that organic chemistry has run rampant through the solar system and beyond. Mars has two small satellites, Phobos and Deimos, which, because of their dark color, seem to be made of (or at least covered by) organic matter. They are widely thought to be captured asteroids from farther out in the solar system. Indeed, there seems to be a vast population of small worlds covered with organic matter: the so-called C- and D-type asteroids in the main asteroid belt between Jupiter and Mars; the nuclei of comets such as Halley's Comet; and the newly discovered class of asteroids near the outermost planets. In 1986 the European Space Agency's *Giotto* spacecraft flew directly into the cloud of dust surrounding Halley's Comet, revealing that its nucleus may be made of as much as 25 percent organic matter.

A fairly abundant type of meteorite on the earth, known as carbonaceous chondrite, is thought to consist of fragments from C-type asteroids in the main belt. Carbonaceous meteorites contain an organic residue rich in aromatic and other hydrocarbons. Scientists have also identified a number of amino acids (the building blocks of the proteins) and nucleotide bases (the "rungs" of the DNA double helix, which spell out the genetic code).

Asteroidal and cometary fragments plunging into the atmosphere of the early earth carried with them vast stores of organic molecules. Some of these survived the intense heating on entry and therefore may have made a significant material contribution to the origin of life. Impacts would have delivered sim-

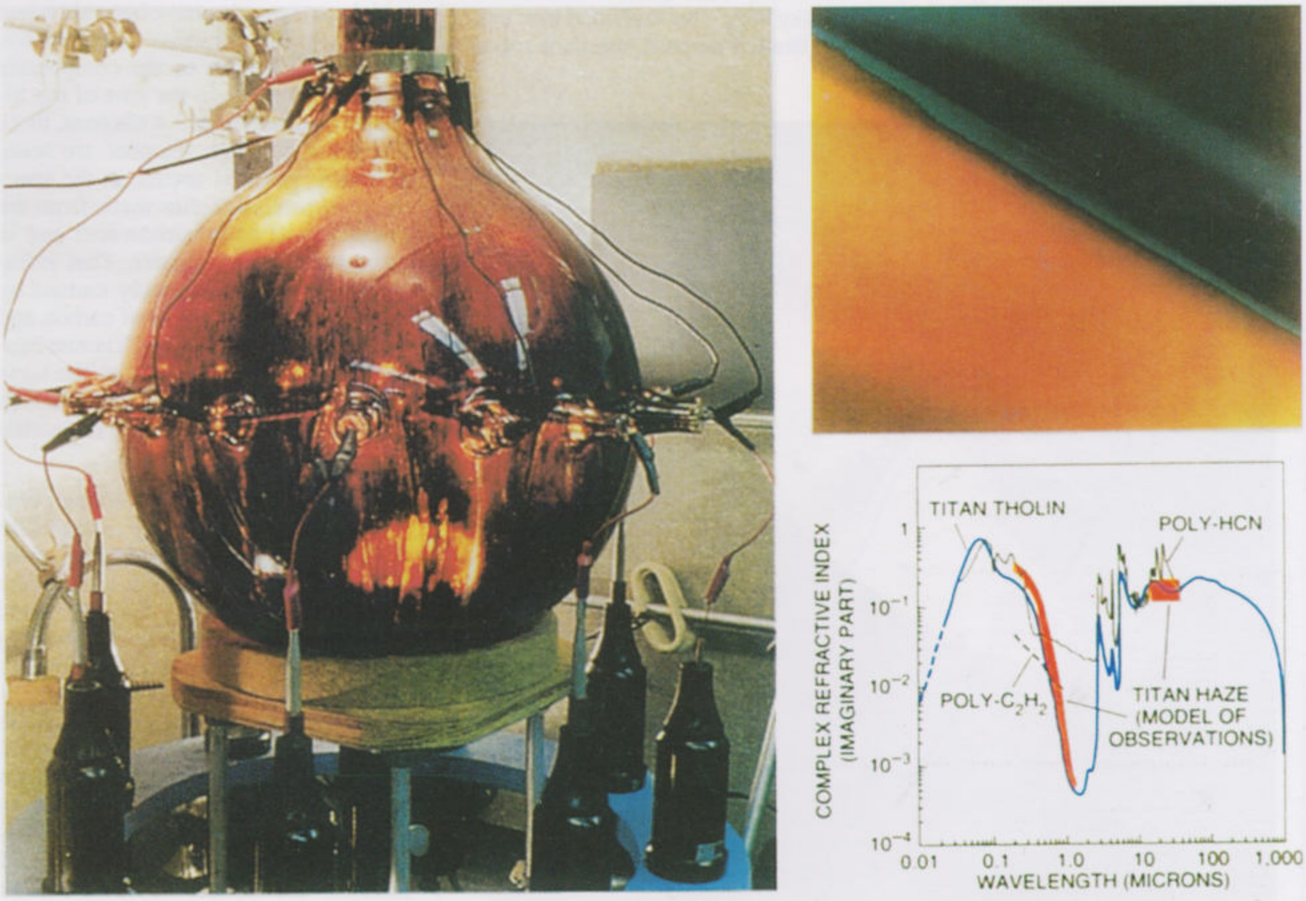
ilar supplies of organic matter, along with water, to other worlds. Those worlds need not be as richly endowed with liquid water as is the earth for critical steps in prebiological chemistry to occur. The water could be found in ponds, in subsurface reservoirs, as thin films on mineral grains or as ice melts formed by impacts.

One of the most fascinating and instructive worlds illustrating prebiological organic chemistry is Saturn's giant moon, Titan (which is as large as the planet Mercury). Here we can see the synthesis of complex organic molecules happening before our eyes. Titan has an atmosphere 10 times as massive as the earth's, composed mainly of molecular nitrogen, along with a few percent to 10 percent methane. When *Voyager 2* approached Titan in 1981, it could not see the surface, because this world is entirely socked in by an opaque, reddish orange haze. The surface temperature is very low, about 94 kelvins, or -179 degrees Celsius. If we can judge from its density (much lower than that of solid rock) and from the composition of nearby worlds, Titan should have a great deal of water ice on and near its surface. A few simple organic molecules—hydrocarbons and nitriles—are found to be minor constituents of Titan's atmosphere.

Ultraviolet light from the sun, charged particles trapped in Saturn's magnetosphere and cosmic rays all bombard Titan's atmosphere and initiate chemical reactions there. When W. Reid Thompson of Cornell University and I considered the effects of ultraviolet irradiation and simulated those of auroral electron bombardment, we found the results agree well with the observed abundances of gaseous organic constituents.

My colleague Bishun N. Khare and I at Cornell simulated the pressure and composition of the appropriate levels in Titan's atmosphere and irradiated the gases with charged particles. The experiment produced a dark, organic solid that we call Titan tholin from the Greek word for "muddy." When we measure the optical constants of Titan tholin, we find that it beautifully matches the optical constants derived from observations of the Titan haze. No other proposed material comes close.

Organic molecules continually form in the upper atmosphere of Titan and slowly fall out as new tholins are generated in the upper air. If this process has continued over the past four billion years Titan's surface must be covered by tens, maybe even hundreds, of meters of tholin and other organic products. Moreover, Thompson and I have calculated that over the history of the solar system, a typical location on Titan has something like a 50-50 chance of having experienced centuries of liquid water from the heat released by impacts. When we mix Titan tholin with water in the laboratory, we make amino acids. There are also traces of nucleotide bases, polycyclic aromatic hydrocarbons and a wonderful brew of other compounds. If 100 million years is enough for the origin of life on the earth, could 1,000 years be enough for



LABORATORY SIMULATION of Titan's nitrogen-methane atmosphere (*left*) yields a tarlike accumulation of complex organic molecules, which the author calls Titan tholin. Analogous chemical reactions may give rise to the haze that obscures Titan's surface (*top right*). The optical characteristics of Titan tholin closely match those of Titan's haze (*bottom right*). When combined with liquid water, Titan tholin produces amino acids, nucleotide bases and other molecules important to terrestrial life. Such molecules might have formed in temporary lakes created by cometary impacts on Titan.

it on Titan? Could life have started on Titan during the centuries following an impact, when lakes of water or water-ice slurries briefly formed? The first close-up examination of Titan—by a Saturn orbiter and Titan entry probe—is scheduled to occur when the ESA-NASA Cassini mission reaches the Saturn system in about 2004.

When we look beyond our solar system, into the gas and grains that populate interstellar space, again we find striking signs of the prevalence of organic chemistry. Astronomers examining microwaves emitted and absorbed by molecules at distinctive frequencies have identified more than four dozen simple organic compounds in interstellar space—hydrocarbons, amines, alcohols and nitriles, some of them having long, straight carbon chains, such as $HC_{11}N$. When a cloud of interstellar dust grains lies between the earth and some more distant infrared source, it is possible to determine which infrared wavelengths are absorbed by the grains and hence to learn about their composition.

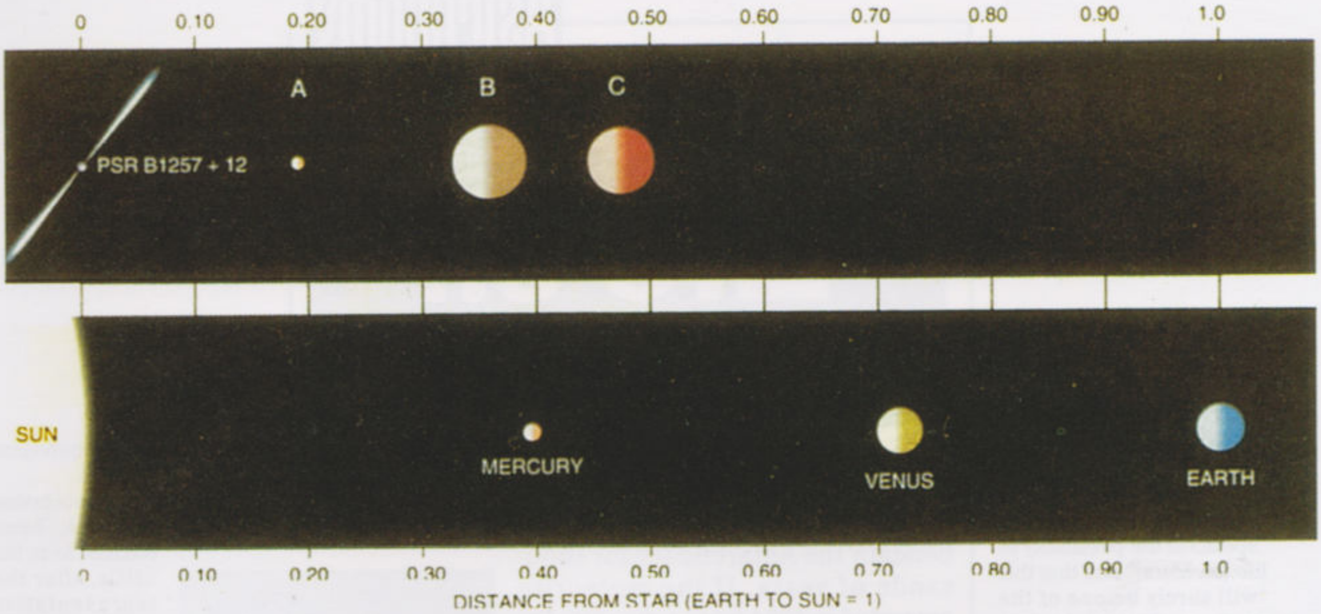
Some of the missing infrared light is widely presumed to have been absorbed by polycyclic aromatics, complex hydrocarbons similar to the compounds found in coal tar. In the part of the infrared spectrum near 3.4 microns, three distinct absorption features are seen. The same patterns appear in the spectra of comets, in tholins made from the irradiation of hydrocarbon ices and in meteoritic organic matter. That infrared fingerprint is probably caused by linked (aliphatic) groups of carbon and hydrogen: $-CH_3$ and $-CH_2$. Yvonne Pendleton and her colleagues at the NASA Ames Research Center find that the best spectral fit seems to be with meteoritic organic matter.

The infrared match among comets, asteroids and interstellar clouds may represent the first direct evidence that asteroids and comets contain organic matter that originated on interstellar grains before gathering together in the infant solar system. But the data are also amenable to an opposite interpretation—that some of the organic matter that formed in the early solar nebula accumulated into asteroids and comets, while some was ejected by the sun into interstellar space. If 100 billion other stars did likewise, they could account for a significant fraction of the organic matter in all the interstellar grains in the galaxy. The prevalence of organic material in the outer solar system, in comets that come from far beyond the outermost planets and in interstellar gas and grains strongly suggests that complex organic matter—

Does Intelligent Life Exist on Other Worlds?

The search for extraterrestrial intelligence is an attempt to use large radio telescopes, sophisticated receivers and modern data analysis to detect hypothetical signals sent our way by advanced civilizations on planets around other stars. Necessarily, there are great uncertainties in selecting the appropriate wavelength, band pass, polarization, time constant and decoding algorithm with which to search for those signals. Nevertheless, radio technology is inexpensive, likely to be discovered early in the evolution of a technological civilization, readily detectable (not just over interplanetary distances, as *Galileo* has done, but over vast interstellar distances) and capable of transmitting enormous amounts of information. The first large-scale, systematic search program, covering a significant fraction of the wavelengths thought optimal for interstellar communication, was initiated by the National Aeronautics and Space Administration on October 12, 1992. Congress canceled the program a year later, but it will soon be resuscitated using private money. Meanwhile some smaller efforts have made provocative findings.

One promising project is the Megachannel Extraterrestrial Array (META), which is led by Paul Horowitz, a physics professor at Harvard University, and funded mainly by the Planetary Society, the largest space interest group in the world. The antenna used for META appears below. After five years of continuous sky survey and two years of follow-up, Horowitz and I found a handful of candidate radio signals that have extremely narrow bandwidths, that do not seem to share the earth's rotation and that cannot be attributed to specific sources of noise or interference. The only trouble is that none of these sources repeats, and in science nonrepeating data are usually not worth much. The tantalizing aspect of the META findings is that the five strongest signals all lie in the plane of the Milky Way. The likelihood that this alignment happens by chance is something like 0.5 percent. We think more comprehensive searches are worth doing.



EXTRASOLAR PLANETS seem to orbit the star PSR B1257+12, the tiny, dense remnant of an ancient supernova explosion. The spacing of the three surrounding planets—known as A, B and C—resembles the circumstances of our solar system (the sizes of the planets are not drawn to scale). It is possible that a more distant, habitable planet circles this stellar corpse. There is also increasing evidence that planetary systems orbit many sunlike stars, which offer better prospects for life.

relevant to the origin of life—is widely spread throughout the Milky Way.

Organic molecules on bone-dry interstellar grains fried by ultraviolet light and cosmic rays seem an unlikely habitat for the origin of life, however. Life seems to need liquid water, which in turn seems to require planets. Astronomical observations increasingly indicate that planetary systems are common. A surprisingly large number of nearby young stars of roughly solar mass are surrounded by just the kind of disks of gas and dust that scientists going back to Immanuel Kant and Pierre Simon, the Marquis de Laplace, say is needed to explain the origin of the planets in our system. These disks provide a persuasive though still indirect indication that there is a multitude of planets, presumably including earthlike worlds, around other stars.

George W. Wetherill of the Carnegie institution of Washington has developed detailed models for predicting the distribution of the planets that should be formed in such circumstellar disks. Meanwhile James F. Kasting of Pennsylvania State University has calculated the range of distances from their suns at which planets can support liquid water on their surfaces. Taken together, these two lines of inquiry suggest that a typical planetary system should contain one and maybe even two earthlike planets circling at a distance where liquid water is possible.

Recently Alexander Wolszczan, also at Pennsylvania State, unambiguously detected earthlike planets in a place where most astronomers least expected to find them: around a pulsar, the swiftly spinning neutron-star remnant from a supernova explosion. Based on variations in the timing of radio emissions from the pulsar PSR B1257+12, Wolszczan has deduced the presence of three planets (so far called only A, B and C) orbiting the pulsar.

These worlds are closer to their star than the earth is to ours, and PSR B1257+12 emits in charged particles several times as much energy as does the sun in electromagnetic radiation. If all the charged particles intercepted by A, B and C are transformed into heat, these worlds must almost certainly be too hot for life. But Wolszczan finds hints of at least one additional planet situated farther from the pulsar. For all we know, this superficially unpromising system, 1,400 light-years from the earth, may contain a dark but habitable planet. It is not clear whether these planets survived from before the supernova explosion or, more likely, formed afterward from surrounding debris. Either way, their presence suggests that planetary formation is an unexpectedly common and widespread process.

Numerous searches for planets in infant and mature sunlike systems are under way. The pace of exploration is becoming so quick, and so many new techniques are about to be employed, that it seems likely that in the next few decades astronomers will begin accumulating a sizable inventory of planets around nearby stars.

We have every reason to believe that there are many water-rich worlds something like our own, each provided with a generous complement of complex organic molecules. Those planets that circle sunlike stars could offer environments in which life would have billions of years to arise and evolve. Should not there be an immense number and diversity of inhabited worlds in the Milky Way? Scientists differ about the strength of this argument, but even at its best it is very different from actually detecting life elsewhere. That monumental discovery remains to be made.

FURTHER READING

MARS. Edited by H. H. Kieffer, B. M. Jakosky, C. Snyder and M. S. Matthews. University of Arizona Press, 1992.

TITAN: A LABORATORY FOR PREBIOLOGICAL ORGANIC CHEMISTRY. Carl Sagan, W. Reid Thompson and Bishun N. Khare in *Accounts of Chemical Research*, Vol. 25, No. 7, pages 286-292; July 1992.

FIVE YEARS OF PROJECT META: AN ALL-SKY NARROW-BAND RADIO SEARCH FOR EXTRATERRESTRIAL SIGNALS. Paul Horowitz and Carl Sagan in *Astrophysical Journal*, Vol. 415, No. 1, pages 218-233; September 20, 1993.

A SEARCH FOR LIFE ON EARTH FROM THE GALILEO SPACECRAFT. Carl Sagan et al. in *Nature*, Vol. 365, No. 6448, pages 715-721; October 21, 1993.

Article 7

Astronomy

Focus

Life on Mars?

When scientists launched the Viking probes to Mars in search of life, little did they know that the answers they were seeking may have laid deep beneath the Antarctic ice for thousands of years. If the meteorite extracted in 1984 from a desolate ice field at the south pole really does hold evidence of primitive life on Mars, it will answer the question: are we really alone in the Universe? And if life began independently on two planets of our solar system, how many other planets in our galaxy and beyond could be harbouring life? Although there was great excitement surrounding the announcement that a team of scientists from NASA and Stanford University have found what they believed to be evidence of early life on Mars, their results are far from conclusive. In this special Focus, Monica Grady and Ian Wright, who have studied the meteorite in question, review the background to the announcement and explain just what the US team claims to have found. Mike Lancaster of the University of London Observatory looks millions of years back in time when Mars may have been a warmer, wetter place and more hospitable to life. The Martian canals turned out to be an optical illusion, but space probes have revealed a surface criss-crossed by channels cut by flowing water. Where did that water go and could it support life? Finally, Stephen Baxter looks at the way science fiction has treated the subject of life on Mars.

On August 7, 1996, the world was stunned to hear President Clinton announce that a meteorite from Mars "speaks of the possibility of life [on Mars]" and that this "will surely be one of the most stunning insights into our universe that science has ever uncovered". This was followed by a cautiously optimistic Dan Goldin, the NASA administrator, announcing that a team of US scientists had found "compelling but not conclusive" evidence for life on Mars.

The scientists, from NASA's Johnson Space Center (JSC) and Stanford University, had analysed a Martian meteorite (Allan Hills 84001, or ALH 84001) and made several observations, which, taken together, allowed them to conclude that they were seeing the fossilised remains of primitive organisms. The findings were published in the journal Science the following week.

What is the background to this remarkable announcement? And if true, what implications does it have for future planetary exploration?

What are Martian meteorites?

There are 12 meteorites in our collections that in all probability come from Mars. Of course we cannot be absolutely certain of this proposition until a rock is returned directly from the planet to Earth *via* a space mission. However, it should be realised that for the scientists that study these Martian meteorites there is little doubt of their origin – the time for debating this issue was effectively ended in the early 1980s with the recognition of meteorites that have a lunar origin.

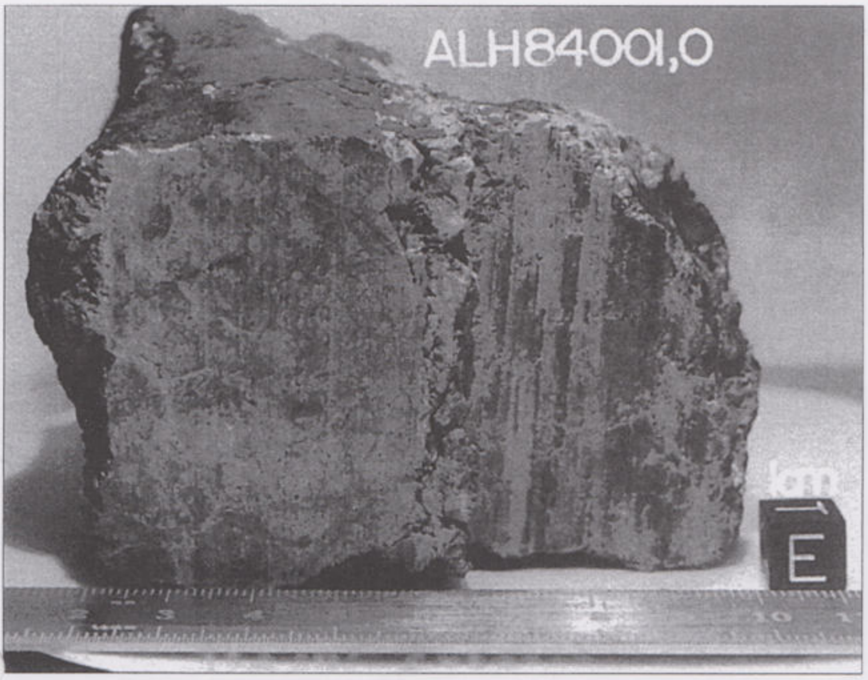
Of the 12 Martian meteorites, six have been found in Antarctica, not because there is a higher incidence of their falling at the Earth's poles, but because there is a greater chance of them being collected there. This is because the icy desert preserves the stones for thousands (sometimes millions) of years – elsewhere on Earth the Martian rocks, which after all are of a planetary nature, are more difficult to distinguish from equivalent terrestrial materials.

The Martian meteorites have, in the past, been referred to as the SNCs, after the representative members of each of the three subgroups into which they were originally subdivided: Shergotty, Nakhla and Chassigny. When ALH 84001 was classified as Martian, it was also recognised as a fourth sub-group of the Martian family. All the Martian meteorites are igneous rocks – they have solidified from magma (molten rock) at or below Mars' surface. The different groups represent crystallisation at different depths; on Earth, geologists would have labelled them lherzolite, pyroxenite, dunite or basalt, etc. Some of the rocks have been altered by fluids, others appear to be dry. Many of the twelve are shocked, but all have one thing in common: they come from the same parent planet.

How do we know they are from Mars?

How can we be so sure the parent planet is Mars? What sets these rocks apart from other meteorites? There are several features that point to a Martian origin, but to appreciate them, we must first understand something of the nature of "regular" meteorites, those that are fragments from the Asteroid Belt found between the orbits of Mars and Jupiter.

The Solar System formed about 4,550 million years ago from a turbulent cloud of gas, ices and dust. The dust aggregated into increasingly larger bodies, forming the inner, rocky planets, whilst gas and ices accumulated in the giant planets Jupiter,



This 4.5 billion year old rock, labelled meteorite ALH84001, is believed to have once been a part of Mars and contains fossil evidence that primitive life may have existed on Mars more than 3.6 billion years ago. Photo: NASA.

Saturn, Uranus and Neptune (as well as Kuiper Belt and Oort cloud objects). The asteroid belt marks the boundary between the inner and outer planets. The bodies of the asteroid belt, also known as minor planets were prevented by Jupiter's gravity from aggregating into a single planetary body. So, as the Solar System settled into something approaching the geometry we have today, the asteroids cooled and solidified. The date of their crystallisation, the formation age, is thus about 4,550 million years, the age at which the Solar System formed. This is the age of all "regular" meteorites from the Asteroid Belt and, indeed, the age of the Sun, and planets like Earth and Mars.

Not all the bodies in the Solar System have completely solidified – the Earth is a prime example, and we have almost daily displays of active volcanoes erupting molten magma. The solidified rock from a contemporary eruption has a formation age of zero. Igneous rocks from different periods in Earth's history have different ages, but all are younger than 3,800 million years – the original rocks that formed the Earth have all been recycled through geological processes.

Eleven of the twelve Martian meteorites have young crystallisation ages, between 180 million years and 1,300 million years. In other words, they have come from a body that supported molten rocks as recently as 180 million years ago – i.e. not the Asteroid Belt. There are few rocky bodies in the Solar System to which this applies: Venus, Earth, Moon , Mars, and some of the satellites of the giant planets.

It is dynamically difficult (although not impossible) to project rocks from Venus to Earth – however, the thick Venusian atmosphere would cause small bodies of ejected rock to melt and vaporise through frictional heating. It is unlikely that ejecta from the satellites of Jupiter or Saturn would escape the gravitational attraction of the parent planets. So, by a process of elimination, Earth, Moon and Mars are left as prime candidates for the origin of these meteorites. Earth and Moon can be ruled out on the grounds of the chemistry of the rocks: the Martian meteorites all have similar oxygen isotopic compositions to each other, but different from the Earth, or the Moon.

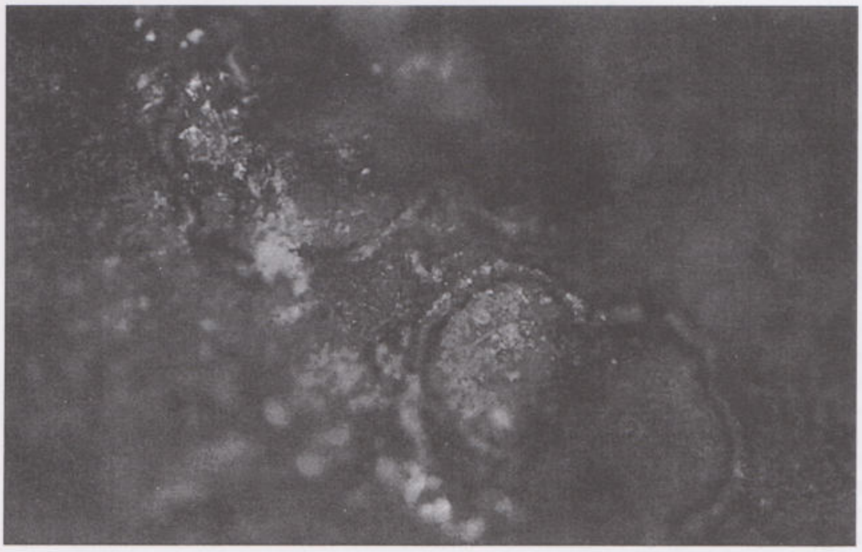
Oxygen isotopic composition (the relative amounts of the three stable isotopes of oxygen ^{16}O, ^{17}O and ^{18}O present in a rock) is a characteristic signature of a parent body: samples with similar isotopic compositions belong to the same family and presumably come from the same parent. Since the oxygen isotopic composition of the Martian Meteorites is different from that of the Earth, they cannot have come from our home planet, leaving Mars as the most viable proposition for the meteorites' origin.

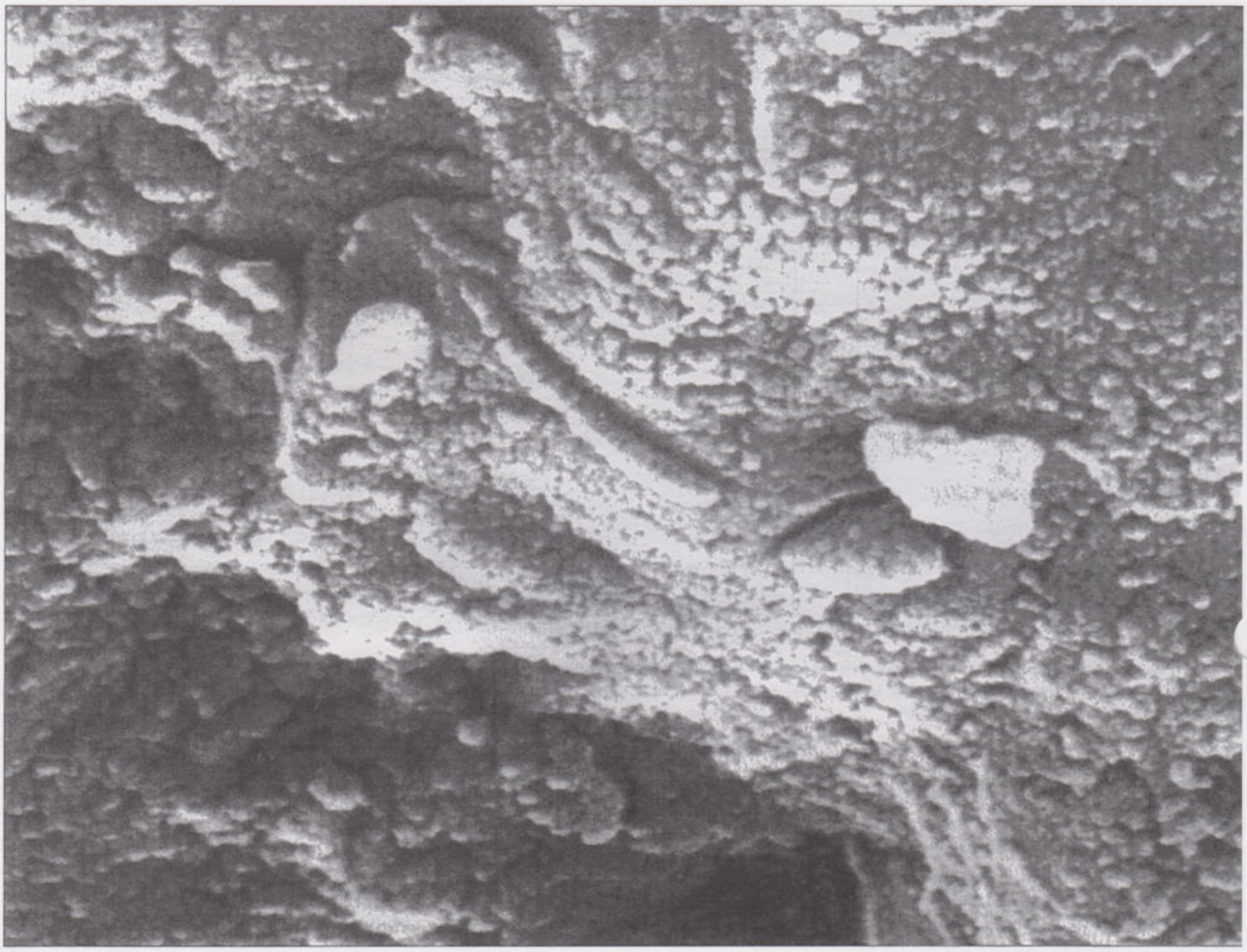
This may all sound rather circumstantial – however, there is one more relevant observation, and this is killer evidence. Recall that the Viking landers of 1976 returned data on the chemical and isotopic composition of Mars' atmosphere. Analysis of one of the Martian meteorites (EET A79001, a basalt from Antarctica collected in 1979) showed the sample to contain, within pockets of shock-produced glass (1-2 cm across), small quantities of Martian atmospheric gases.

Recovery of ALH 84001

In 1984, the Antarctic Search for Meteorites team of US scientists paid its annual visit to Antarctica, to search for meteorites on the Allan Hills ice-field. Among the rocks found that year was a 1.9 kg specimen that, on return to the Antarctic Meteorite Curatorial Facility at JSC in Houston, Texas, was classified as a basalt. Unfortunately, it was thought to be a diogenite, a basalt from the Asteroid Belt, possibly from Asteroid Vesta. Ten years later, a sample of ALH 84001 was requested by a scientist studying diogenites. He immediately realised that the meteorite was unlike any other diogenite, and so sent it for oxygen isotope analysis, when it was found that the oxygen composition of ALH 84001 was close to that of the other Martian mete-

This photograph shows orange-coloured carbonate mineral globules found in the ALH84001 meteorite. These carbonate minerals are believed to have formed more than 3.6 billion years ago. Their structure and chemistry suggest that they may have been formed with the assistance of primitive bacteria-like organisms. Photo: NASA.





This high-resolution scanning electron microscope image shows an unusual tube-like form that is less than 1/100th the width of a human hair in size found in meteorite ALH84001. Although this structure is not part of the research published in *Science*, it is located in a similar carbonate glob in the meteorite. This structure will be the subject of future investigations that could confirm whether or not it is fossil evidence of primitive life on Mars. Photo: NASA.

orites. And so an additional member of the Martian family was recognised. However, ALH 84001 differs from its relatives in several ways: it is old, having crystallised not long after Mars accreted. Although it is an igneous rock, it has suffered extensive alteration, leading to the formation of large patches of bright orange carbonates.

The announcement, around February 1994, that an additional and unusual Martian meteorite was available for study led to a flurry of applications from scientists to acquire fragments of the rock for study. A team of UK scientists, based at the Open University and the Natural History Museum, received samples of material as part of their investigation of carbon, nitrogen and argon in Martian meteorites.

Using the technique of high resolution stepped combustion-mass spectrometry, the scientists discovered that ALH 84001 contained more carbon, in the form of organic compounds, than any other Martian meteorite they had studied. The technique used, however, did not indicate the nature of the organic molecules, and their carbon isotopic composition was not diagnostic – the organics might have been Martian or terrestrial. The carbonates, on the other hand, had a distinctive carbon isotopic composition, indicating that they had been formed from fluids in contact with Mars' atmosphere.

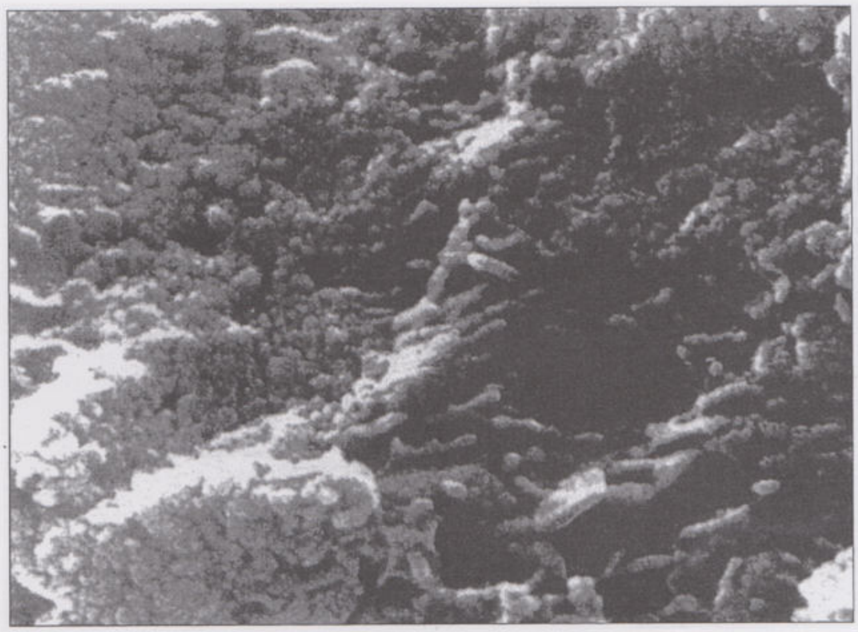
Following this work, US scientists attempted to characterise the nature of the organics present – and it is this work that hit the headlines in August.

Life in ALH 84001?

Dave McKay and his team studied ALH 84001 using a variety of techniques. They used a laser microprobe to release organic molecules from within the meteorite, followed by analysis of the compounds using mass spectrometry. This showed that most of the organics were closely associated with the carbonates (which have an indisputable Martian origin) and so they concentrated their efforts on these materials. Some of the organic compounds are polycyclic aromatic hydrocarbons, or PAHs. Benzene is a good example of a PAH, although the compounds identified in ALH 84001 are more complicated than this. If the carbonates are Martian, argued the authors, then so too must be the organics.

The carbonates were studied by high resolution transmission and scanning electron microscopy (TEM and SEM), the former technique looking at slices through the carbonate, the latter examining the outer surfaces. Results from TEM showed that concentrated in zones within the carbonates were tiny grains of iron oxide (magnetite) and iron sulphides (pyrrhotite and greigite). The individual magnetite grains had shapes not usually observed in meteoritic magnetites: cuboid and teardrop. The association of magnetite and iron sulphides with dissolution hollows in the carbonate grains also suggested to McKay and his team that they were observing a disequilibrium mineral assemblage such as is often characteristic of biogenic processes.

Examination of the outer surface of the carbonate areas by SEM showed that the surface had an irregular texture, not consistent with growth or cleavage surfaces. Also on the surface were oval-shaped and elongated structures. These, by analogy with terrestrial systems were interpreted as the remains of nanobacteria (i.e. bacteria of a few tens of nanometres in size).



This electron microscope image shows tubular structures of likely Martian origin, according to NASA researchers. Photo: NASA.

Other ovoid features occur in the iron-rich rims of the carbonates, and might therefore be associated with the magnetite and sulphide grains, and thus represent the products of microbial activity.

In summarising their findings, McKay's team is very careful to point out that each observation is capable of an alternative, inorganic processes. But some of these processes might require ad hoc explanations of changing conditions, such as alternating oxidising and reducing environments, variable acidity of the fluids, etc. In the researchers' opinion, the most likely explanation that satisfies all the observations is that the carbonate patches, PAHs, the magnetite and the microstructures "could be the fossil remains of a past Martian biota".

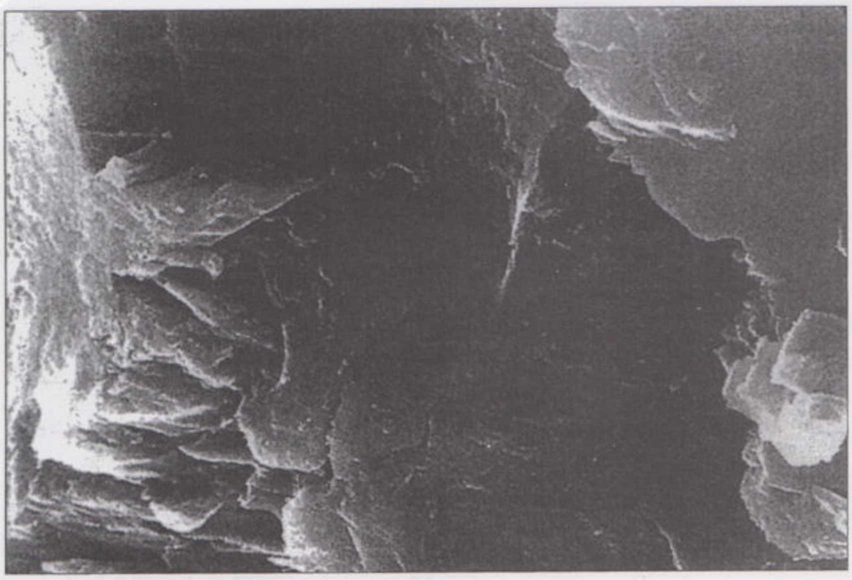
This opinion is not universally accepted. At the press conference held to announce the findings, Professor William Schopf, a microfossil expert, pointed out several flaws in the arguments. To start with, the magnetite occurs as discrete grains, and not as a sequence of several connected grains, such as might be found in a terrestrial organism. The grains are also about 100 times smaller than those characteristic of biogenic processes on Earth. They do not appear to contain the voids or cavities observed in terrestrial microfossils that are a sign that fluids associated with life have been present. And structures that could be interpreted as cell walls have not been identified.

In addition to flaws with the microstructure argument, the identification of PAHs is also not a good biomarker. Although PAHs are abundant in coal and petroleum deposits they are not a primary biological product, but are a result of secondary diagenesis (heating that accompanies the burial of rocks, which in turn acts to convert pre-existing organic compounds into PAHs). The rock ALH 84001 gives little indication that such processes have occurred on its parent.

So, given the evidence presented by McKay and his team, and the observations of other commentators, have the remnants of life on Mars been found? On balance, we think not (but then again, scientists are naturally cautious!) – the microfossil evidence is inconclusive, and might be explicable by inorganic processes. But this does not mean that we rule out the possibility that life once existed on Mars. The building blocks of life were undoubtedly present on the planet in the past. In previous times it had a thicker atmosphere, allowing water to flow on its surface: we can still see the dried up remnants of water channels on satellite images. However, when Mars lost its atmosphere, the surface waters also disappeared. Mars is now a dry planet, its surface sterilised by the sun's ultraviolet radiation. However, we do not know what lurks below the surface – could life still exist there today?

One immediate outcome of the announcement of finding life on Mars has been a pledge from President Clinton in support of future planetary exploration programmes, not just to Mars, but the search for planetary systems around other stars. There are six missions to Mars planned in the next three years, none of which is a sample return mission. However, NASA's funding must surely be a little more secure, allowing active planning of a Martian sample return mission, hopefully early in the next millennium, around 2005, or even 2003.

In the centre of this electron microscope image of a small chip ALH84001 are several tiny structures that are possible fossils of primitive, bacteria-like organisms that may have lived on Mars more than 3.6 billion years ago. Photo: NASA



Whatever McKay and his team observed in ALH 84001, primitive microfossils or non-biogenic molecules, they have pushed open further the door to exploration of our neighbouring planet and raised public awareness of how close we are to finding signs of life (extinct or dormant) on other bodies within the solar system.

Recent reports from the Galileo probe that the Jovian satellite Europa might be rich in ice, and parts of Io covered in slush reinforce the claims that life might exist in other corners of our solar system – although the images of microstructures in ALH 84001 remind us that any possible life forms are likely to be very primitive. We might not be alone... but we don't yet have anyone to talk to!

Monica Grady is a research specialist in meteorites in the Department of Mineralogy at the Natural History Museum, and Ian Wright is a Research Fellow in the Planetary Sciences Research Institute at the Open University.

THE SEARCH FOR EXTRATE

Scientific Quest or

A DEBATE BETWEEN ERNST MAYR AND CARL SAGAN

Ernst Mayr –"Darwin's current bulldog" is how Scientific American describes Ernst Mayr, recalling Thomas Huxley, the 19th-century scientist who is remembered largely because of his vociferous defense of the theory of evolution. Mayr will be remembered for championing evolution as well as for his many achievements in the field of biology. In systematics, ornithology, evolutionary biology and the history of science, he stands as a giant of the 20th century. He has been awarded the National Medal of Science, the Balzan Prize for his contributions to evolutionary biology and the Sarton Medal for his work in the history of science. As the Alexander Agassiz Professor of Zoology, Emeritus, of Harvard University, the 91-year-old Mayr still works every day, tackling challenging subjects, such as the Search for Extraterrestrial Intelligence.

Since humans first looked up, they have seen in the skies the phantoms of their wondering minds. If there is one thread that links the ancient Greek philosophers to modern space scientists, it is the uncertainty about the plurality of worlds. Vast and ancient beyond ordinary human understanding, the universe leaves us pondering the ultimate significance, if any, of our tiny but exquisite life-bearing blue planet.

With the development of technology and our present understanding of the laws of nature, the human species is now in a position where the possibility of extraterrestrial civilizations can be verified by experiment. But because we have yet to find a single piece of concrete evidence of alien intelligence, a philosophical battle has arisen between those who might be called contact optimists–who generally embrace SETI–and the proponents of the uniqueness hypothesis, which suggests that Earth is the only technical civilization in our galaxy.

In these pages, we present both sides of this philosophical and scientific battle. Which view is more palatable to you? Read on and decide for yourself. –Guillermo A. Lemarchand

This debate between two of the most prominent scientists of the 20th century first appeared in *The Bioastronomy News*, a special-interest newsletter published by The Planetary Society. The debate was conceived by newsletter editor Guillermo A. Lemarchand, a researcher at the Argentine Institute of Radio Astronomy and a leader in the Society's META II project.

THE IMPROBABILITY OF SUCCESS

by Ernst Mayr

What is the chance of success in the Search for Extraterrestrial Intelligence? The answer to this question depends on a series of probabilities. I have attempted to make a detailed analysis of this problem in a German publication (Mayr, 1992) and shall attempt here to present in English the essential findings of this investigation. My methodology consists in asking a series of questions that narrow down the probability of success.

How Probable Is It That Life Exists Somewhere Else in the Universe?

Even most skeptics of the SETI project will answer this question optimistically. Molecules that are necessary for the origin of life, such as amino acids and nucleic acids, have been identified in cosmic dust, together with other macromolecules, and so it would seem quite conceivable that life could originate elsewhere in the universe.

Some of the modern scenarios of the origin of life start out with even simpler molecules—a beginning that makes an independent origin of life even more probable. Such an independent origin of life, however, would presumably result in living entities that are drastically different from life on Earth.

RRESTRIAL INTELLIGENCE:
Hopeful Folly?

Carl Sagan - Carl Sagan is the David Duncan Professor of Astronomy and Space Science and Director of the Laboratory for Planetary Studies at Cornell University; Distinguished Visiting Scientist at the Jet Propulsion Laboratory of the California Institute of Technology; and cofounder and President of The Planetary Society. He is one of the few astronomers with a background in biology: research assistant to Nobel laureate geneticist H.J. Muller at Indiana University, Visiting Assistant Professor of Genetics at the Stanford University Medical School and author of the article "Life" in the Encyclopaedia Britannica. He is coauthor, with I.S. Shklovskii, of the classic SETI text, Intelligent Life in the Universe, and has made many research contributions to the study of the prebiological organic chemistry on Earth and in the outer solar system and of the origin of life.

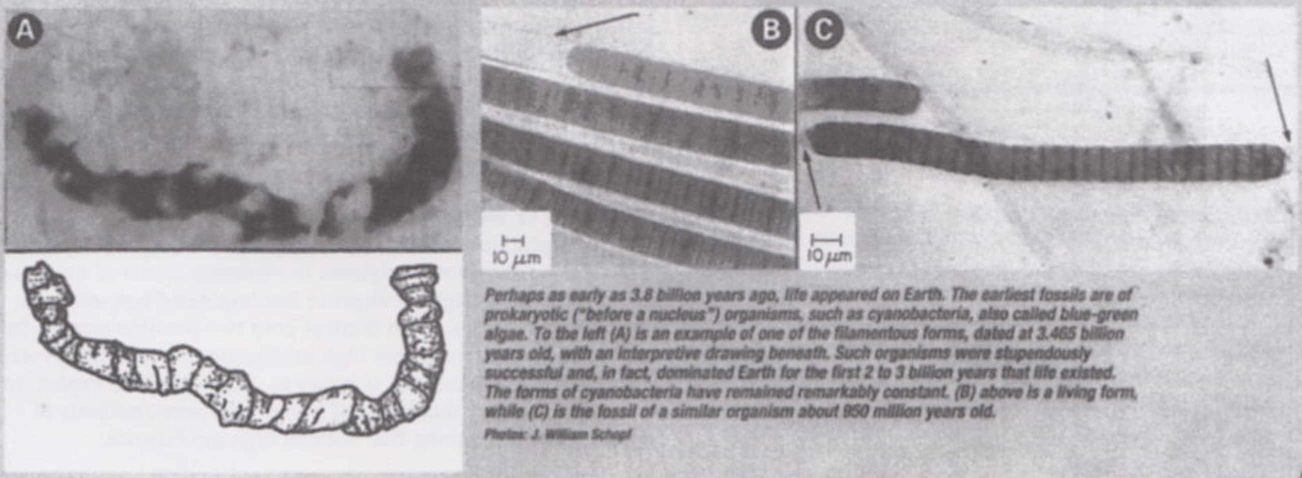
Where Can One Expect to Find Such Life?

Obviously, only on planets. Even though we have up to now secure knowledge only of the nine planets of our solar system, there is no reason to doubt that in all galaxies there must be millions if not billions of planets. The exact figure, for instance, for our own galaxy can only be guessed.

How Many of These Planets Would Have Been Suitable for the Origin of Life?

There are evidently rather narrow constraints for the possibility of the origin and maintenance of life on a planet. There has to be a favorable average temperature; the seasonal variation should not be too extreme; the planet must have a suitable distance from its sun; it must have the appropriate mass so that its gravity can hold an atmosphere; this atmosphere must have the right chemical composition to support early life; it must have the necessary consistency to protect the new life against ultraviolet and other harmful radiations; and there must be water on such a planet. In other words, all environmental conditions must be suitable for the origin and maintenance of life.

One of the nine planets of our solar system had the right



Perhaps as early as 3.8 billion years ago, life appeared on Earth. The earliest fossils are of prokaryotic ("before a nucleus") organisms, such as cyanobacteria, also called blue-green algae. To the left (A) is an example of one of the filamentous forms, dated at 3.465 billion years old, with an interpretive drawing beneath. Such organisms were stupendously successful and, in fact, dominated Earth for the first 2 to 3 billion years that life existed. The forms of cyanobacteria have remained remarkably constant. (B) above is a living form, while (C) is the fossil of a similar organism about 950 million years old.

Photos: J. William Schopf

kind of mixture of these factors. This, surely, was a matter of chance. What fraction of planets in other solar systems will have an equally suitable combination of environmental factors'? Would it be one in 10, or one in 100, or one in 1,000,000? Which figure you choose depends on your optimism. It is always difficult to extrapolate from a single instance. This figure, however, is of some importance when you are dealing with the limited number of planets that can be reached by any of the SETI projects.

What Percentage of Planets on Which Life Has Originated Will Produce Intelligent Life?

Physicists, on the whole, will give a different answer to this question than biologists. Physicists still tend to think more deterministically than biologists. They tend to say that if life has originated somewhere, it will also develop intelligence in due time. The biologist, on the other hand, is impressed by the improbability of such a development.

Conditions According to Mayr That Must Be Met for the Success of SETI

	Condition	Likelihood
1	Extraterrestrial life must be able to originate repeatedly.	Probable
2	Other habitable planets similar to Earth must be available.	Probable
3	The planet must have conditions that enable the development of intelligent life.	Improbable
4	Extraterrestrial life must adapt "toward" high intelligence.	Highly improbable
5	The extraterrestrial life-forms must grow intelligent enough to found a civilization.	Improbable
6	The extraterrestrial civilization must be able to send and receive signals.	Improbable
7	The life-forms' sense organs must be adapted to receive electronic signals.	Improbable
8	The civilization must be long-lived and sending signals for a long time.	Improbable

Life originated on Earth about 3.8 billion years ago, but high intelligence did not develop until about half a million years ago. If Earth had been temporarily cooled down or heated up too much during these 3.8 billion years, intelligence would have never originated.

When answering this question, one must be aware of the fact that evolution never moves in a straight line toward an objective ("intelligence"), as happens during a chemical process or as a result of a law of physics. Evolutionary pathways are highly complex and resemble more a tree with all of its branches and twigs.

After the origin of life—that is, 3.8 billion years ago—life on Earth consisted for 2 billion years only of simple prokaryotes, cells without an organized nucleus. These bacteria and their relatives developed surely 50 to 100 different (some perhaps very different) lineages, but, in this enormously long time, none of them led to intelligence. Owing to an astonishing, unique event that is even today only partially explained, about 1,800 million years ago the first eukaryote originated, a creature with a well-organized nucleus and the other characteristics of "higher" organisms. From the rich world of the protists (consisting of only a single cell), there eventually originated three groups of multicellular organisms: fungi, plants and animals. But none of the millions of species of fungi and plants was able to produce intelligence.

The animals (Metazoa) branched out in the Precambrian and Cambrian time periods to about 60 to 80 lineages (phyla). Only a single one of them, that of the chordates, led eventually to genuine intelligence. The chordates are an old and well-diversified group, but only one of its numerous lineages, that of the vertebrates, eventually produced intelligence. Among the vertebrates, a whole series of groups evolved—types of fishes, amphibians, reptiles, birds and mammals. Again, only a single lineage, that of the mammals, led to high intelligence. The mammals had a long evolutionary history which began in the Triassic Period, more than 200 million years ago, but only in the latter part of the Tertiary Period—that is, some 15 to 20 million years ago—did higher intelligence originate in one of the circa 24 orders of mammals.

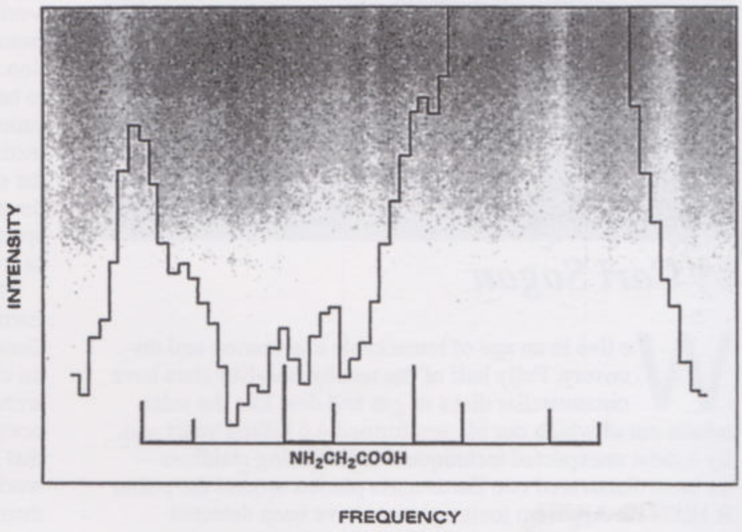
The elaboration of the brain of the hominids began less than 3 million years ago, and that of the cortex of *Homo sapiens* occurred only about 300,000 years ago. Nothing demonstrates the improbability of the origin of high intelligence better than the millions of phyletic lineages that failed to achieve it.

How many species have existed since the origin of life? This figure is as much a matter of speculation as the number of planets in our galaxy. But if there are 30 million living species, and if the average life expectancy of a species is about 100,000 years, then one can postulate that there have been billions, perhaps as many as 50 billion species since the origin of life. Only one of these achieved the kind of intelligence needed to establish a civilization.

To provide exact figures is difficult because the range of variation both in the origination of species and in their life expectancy is so enormous. The widespread, populous species of long geological duration (millions of years), usually encountered by the paleontologist, are probably exceptional rather than typical.

Why Is High Intelligence So Rare?

Adaptations that are favored by selection, such as eyes or bioluminescence, originate in evolution scores of times independently. High intelligence has originated only once, in human beings. I can think of only two possible reasons for this rarity. One is that high intelligence is not at all favored by natural selection, contrary to what we would expect. In fact, all the other kinds of living organisms, millions of species, get along fine without high intelligence.



The building blocks of life may even be found in interstellar space. This spectrum of a molecular cloud near our galactic center shows peaks at some of the frequencies expected for the glycine molecule (NH_2CH_2COOH), one of the amino acids upon which earthly life is based.

Chart: Lewis E. Snyder; redrawn by B.S. Smith

The other possible reason for the rarity of intelligence is that it is extraordinarily difficult to acquire. Some grade of intelligence is found only among warm-blooded animals (birds and mammals), not surprisingly so because brains have extremely high energy requirements. But it is still a very big step from "some intelligence" to "high intelligence."

The hominid lineage separated from the chimpanzee lineage about 5 million years ago. but the big brain of modern man was acquired less than 300,000 years ago. As one scientist has suggested (Stanley, 1992), it required complete emancipation from arboreal life to make the arms of the mothers available to carry the helpless babies during the final stages of brain growth. Thus, a large brain, permitting high intelligence, developed in less than the last 6 percent of the life on the hominid line. It seems that it requires a complex combination of rare, favorable circumstances to produce high intelligence (Mayr, 1994).

How Much Intelligence Is Necessary to Produce a Civilization?

As stated, rudiments of intelligence are found already among birds (ravens, parrots) and among non-hominid mammals (carnivores, porpoises, monkeys, apes and so forth), but none of these instances of intelligence has been sufficient to found a civilization.

Is Every Civilization Able to Send Signals Into Space and to Receive Them?

The answer quite clearly is no. In the last 10,000 years, there have been at least 20 civilizations on Earth, from the Indus, the Sumerian and other Near Eastern civilizations, to Egypt, Greece and the whole series of European civilizations, to the Mayas, Aztecs and Incas, and to the various Chinese and Indian civilizations. Only one of these reached a level of technology that has enabled it to send signals into space and to receive them.

Would the Sense Organs of Extraterrestrial Beings Be Adapted to Receive Our Electronic Signals?

This is by no means certain. Even on Earth, many groups of animals are specialized for olfactory or other chemical stimuli and would not react to electronic signals. Neither plants nor fungi are able to receive electronic signals. Even if there were higher organisms on some planet, it would be rather improbable that they would have developed the same sense organs that we have.

How Long Is a Civilization Able to Receive Signals?

All civilizations have only a short duration. I will try to emphasize the importance of this point by telling a little fable.

Let us assume that there were really intelligent beings on another planet in our galaxy. A billion years ago, their astronomers discovered Earth and reached the conclusion that this planet might have the proper conditions to produce intelligence. To test this, they sent signals to Earth for a billion years without ever getting an answer. Finally, in the year 1800 (of our calendar) they decided they would send signals only for another 100 years. By the year 1900, no answer had been received, so they concluded that surely there was no intelligent life on Earth.

This shows that even if there were thousands of civilizations in the universe, the probability of a successful communication would be extremely slight because of the short duration of the "open window."

One must not forget that the range of SETI systems is very limited, reaching only part of our galaxy. The fact that there are a near infinite number of additional galaxies in the universe is irrelevant as far as SETI projects are concerned.

SETI Success: An Improbability of Astronomic Dimensions

What conclusions must we draw from these considerations? No less than six of the eight conditions to be met for SETI success are improbable. When one multiplies these six improbabilities with each other, one reaches an improbability of astronomic dimensions.

Why are there nevertheless still proponents of SETI? When one looks at their qualifications, one finds that they are almost exclusively astronomers, physicists and engineers. They are simply unaware of the fact that the success of any SETI effort is not a matter of physical laws and engineering capabilities but essentially a matter of biological and sociological factors. These, quite obviously, have been entirely left out of the calculations of the possible success of any SETI project.

References

Ernst Mayr, "Lohnt sich die Suche nach extraterrestrischer Intelligenz [Is It Worthwhile to Search for Extraterrestrial Intelligence'?]," *Naturwissenschaftliche Rundschau*, vol. 45, no. 7, 1992, pp. 264–266.

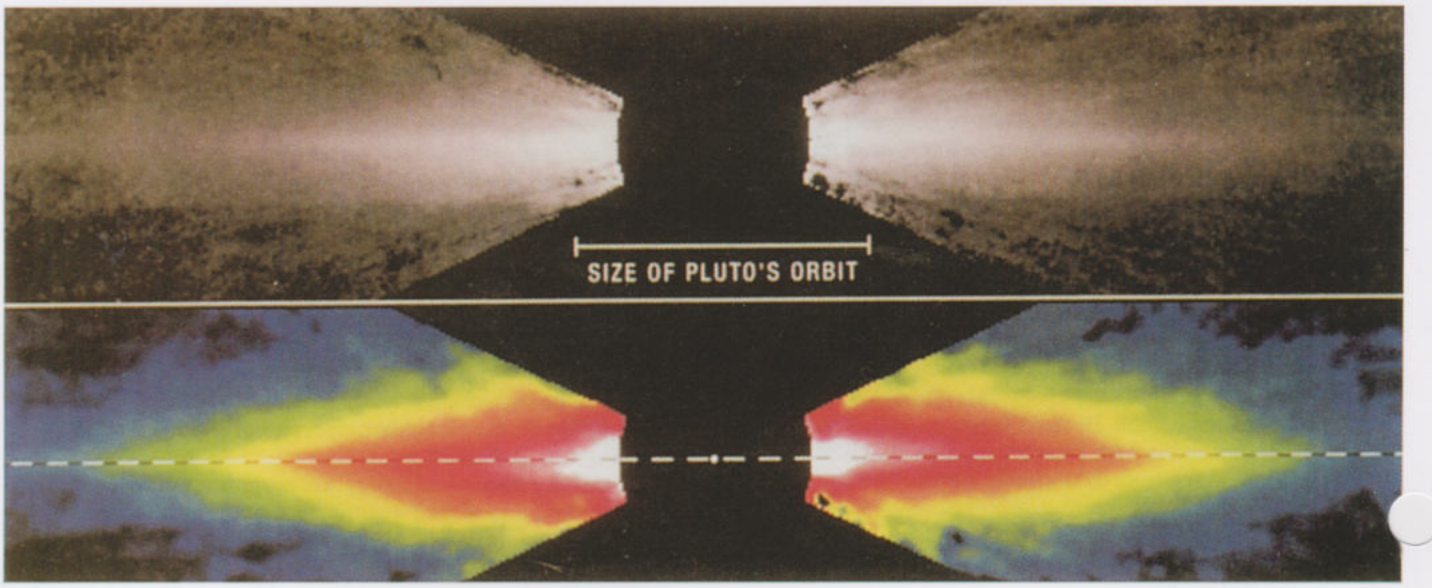
Ernst Mayr, "Does It Pay to Acquire High Intelligence?" *Perspectives in Biology and Medicine*, 1994, pp. 150–154.

S. Stanley, "An Ecological Theory for the Origin of *Homo*," *Paleobiology*, vol. 18, 1992, pp. 237–257.

THE ABUNDANCE OF LIFE-BEARING PLANETS

by Carl Sagan

We live in an age of remarkable exploration and discovery. Fully half of the nearby Sun-like stars have circumstellar disks of gas and dust like the solar nebula out of which our planets formed 4.6 billion years ago. By a most unexpected technique—radio timing residuals—we have discovered two Earth-mass planets around the pulsar B 1257+12. Apparent jovian planets have been detected around the stars 51 Pegasi, 70 Virginis and 47 Ursae Majoris. A range of new Earth-based and spaceborne techniques—including astrometry, spectrophotometry, radial velocity measurements, adaptive optics and interferometry—all seem to be on the verge of being able to detect jovian-type planets, if they exist, around the nearest stars. At least one proposal (The FRESIP [Frequency of Earth-Sized Inner Planets] Project, a spaceborne spectrophotometric system) holds the promise of detecting terrestrial planets more readily than jovian ones. If there is not a sudden cutoff in support, we are likely entering a golden age in the study of the planets of other stars in the Milky Way galaxy.



In 1984, astronomers Brad Smith and Rich Terrile discovered a disk of dust 200 billion kilometers (about 120 billion miles) wide surrounding the star Beta Pictoris. This was exactly what was predicted by theories of planetary formation. The innermost region was clear of dust, as if planets orbiting there had swept it up, but there was little else to suggest that there might be planets orbiting that star—until the Hubble Space Telescope returned the images seen here. They reveal that the inner dust disk is warped. This is indirect evidence that a planet roughly Jupiter's size is orbiting the star at a distance that falls within the range found for planets in our solar system. (The dark region in the center of both images is caused by an occulting disk in the telescope.) *Images: Chris Burrows, ESA and NASA*

Once you have found another planet of Earth-like mass, however, it of course does not follow that it is an Earth-like world. Consider Venus. But there are means by which, even from the vantage point of Earth, we can investigate this question. We can look for the spectral signature of enough water to be consistent with oceans. We can look for oxygen and ozone in the planet's atmosphere. We can seek molecules like methane, in such wild thermodynamic disequilibrium with the oxygen that it can only be produced by life. (In fact, all of these tests for life were successfully performed by the *Galileo* spacecraft in its close approaches to Earth in 1990 and 1992 as it wended its way to Jupiter [Sagan et al., 1993].)

The best current estimates of the number and spacing of Earth-mass planets in newly forming planetary systems (as George Wetherill reported at the first international conference on circumstellar habitable zones [Doyle, 1996]) combined with the best current estimates of the long-term stability of oceans on a variety of planets (as James Kasting reported at that same meeting [Doyle, 1996]) suggest one to two blue worlds around every Sun-like star. Stars much more massive than the Sun are comparatively rare and age quickly. Stars comparatively less massive than the Sun are expected to have Earth-like planets, but the planets that are warm enough for life are probably tidally locked so that one side always faces the local sun. However, winds may redistribute heat from one hemisphere to another on such worlds, and there has been very little work on their potential habitability.

Nevertheless, the bulk of the current evidence suggests a vast number of planets distributed through the Milky Way with abundant liquid water stable over lifetimes of billions of years. Some will be suitable for life—our kind of carbon and water life—for billions of years less than Earth, some for billions of years more. And, of course, the Milky Way is one of an enormous number, perhaps a hundred billion, other galaxies.

Need Intelligence Evolve on an Inhabited World?

We know from lunar cratering statistics, calibrated by returned *Apollo* samples, that Earth was under hellish bombardment by small and large worlds from space until around 4 billion years ago. This pummeling was sufficiently severe to drive entire atmospheres and oceans into space. Earlier, the entire crust of Earth was a magma ocean. Clearly. this was no breeding ground for life.

Yet, shortly thereafter—Mayr adopts the number 3.8 billion years ago—some early organisms arose (according to the fossil evidence). Presumably the origin of life had to have occupied some time before that. As soon as conditions were favorable, life began amazingly fast on our planet. I have used this fact (Sagan, 1974) to argue that the origin of life must be a highly probable circumstance; as soon as conditions permit, up it pops!

Now, I recognize that this is at best a plausibility argument and little more than an extrapolation from a single example. But we are data constrained; it's the best we can do.

Does a similar analysis apply to the evolution of intelligence? Here you have a planet burgeoning with life, profoundly changing the physical environment, generating an oxygen atmosphere 2 billion years ago, going through the elegant diversification that Mayr briefly summarized—and not for almost 4 billion years does anything remotely resembling a echnical civilization emerge.

In the early days of such debates (for example, G. G. Simpson's "The Non-prevalence of Humanoids"), writers argued that an enormous number of individually unlikely steps were required to produce something very like a human being, a "humanoid"; that the chances of such a precise repetition occurring on another planet were nil; and therefore that the chance of extraterrestrial intelligence was nil. But clearly when we're talking about extraterrestrial intelligence, we are not talking—despite Star *Trek—of* humans or humanoids. We are talking about the functional equivalent of humans—say, any creatures able to build and operate radio telescopes. They may live on the land or in the sea or in the air. They may have unimaginable chemistries, shapes, sizes, colors, appendages and opinions. We are not requiring that they follow the particular route that led to the evolution of humans. There may be many different evolutionary pathways, each unlikely, but the sum of the number of pathways to intelligence may nevertheless be quite substantial.

In Mayr's current presentation, there is still an echo of "the non-prevalence of humanoids." But the basic argument is, I think, acceptable to all of us. Evolution is opportunistic and not foresighted. It does not "plan" to develop intelligent life a few billion years into the future. It responds to short-term contingencies. And yet, other things being equal, it is better to be smart than to be stupid, and an overall trend toward intelligence can be perceived in the fossil record. On some worlds, the selection pressure for intelligence may be higher; on others, lower.

If we consider the statistics of one, our own case—and take a typical time from the origin of a planetary system to the development of a technical civilization to be 4.6 billion years—what follows? We would not expect civilizations on different worlds to evolve in lockstep. Some would reach technical intelligence more quickly, some more slowly, and—doubtless—some never. But the Milky Way is filled with second- and third-generation stars (that is, those with heavy elements) as old as 10 billion years.

So let's imagine two curves: The first is the probable timescale to the evolution of technical intelligence. It starts out very low; by a few billion years it may have a noticeable value; by 5 billion years, it's something like 50 percent; by 10 billion years, maybe it's approaching 100 percent. The second curve is the ages of Sun-like stars, some of which are very young—they're being born right now—some of which are as old as the Sun, some of which are 10 billion years old. If we convolve these two curves, we find there's a chance of technical civilizations on planets of stars of many different ages—not much in the very young ones, more and more for the older ones. The most likely case is that we will hear from a civilization considerably more advanced than ours. For each of those technical civilizations, there will have been tens of billions or more other species. The number of unlikely events that had to be concatenated to evolve a technical species is enormous, and perhaps there are members of each of those species who pride themselves on being uniquely intelligent in all the universe.

Need Civilizations Develop the Technology for SETI?

It is perfectly possible to imagine civilizations of poets or (perhaps) Bronze Age warriors who never stumble on James Clerk Maxwell's equations and radio receivers. But they are removed by natural selection. The Earth is surrounded by a population of asteroids and comets, such that occasionally the planet is struck by one large enough to do substantial damage.

From what was probably a single progenitor, life on Earth has adapted and evolved into a bewildering variety of forms, with astonishing behaviors that enable these teeming organisms to survive. Consider the sponge: For most of its life, this animal sits attached to one place, sitting its food from seawater flowing through its many cavities. It is not a lifestyle that requires intelligence. But the sponge does possess one remarkable trait. If a sponge is forced through a fine mesh, breaking it into single ceils or clumps of cells, the cells will reassemble themselves into another sponge.

Photo: Brian Parker, Tom Stack & Associates

The most famous is the K-T event (the massive near-Earth-object impact that occurred at the end of the Cretaceous Period and start of the Tertiary) of 65 million years ago that extinguished the dinosaurs and most other species of life on Earth. But the chance is something like one in 2,000 that a civilization-destroying impact will occur in the next century.

It is already clear that we need elaborate means for detecting and tracking near-Earth objects and the means for their interception and destruction. If we fail to do so, we will simply be destroyed. The Indus Valley, Sumerian, Egyptian, Greek and other civilizations did not have to face this crisis because they did not live long enough. Any long-lived civilization, terrestrial or extraterrestrial, must come to grips with this hazard. Other solar systems will have greater or lesser asteroidal and cometary fluxes, but in almost all cases the dangers should be substantial.

Radiotelemetry, radar monitoring of asteroids, and the entire concept of the electromagnetic spectrum are part and parcel of any early technology needed to deal with such a threat. Thus, any long-lived civilization will be forced by natural selection to develop the technology of SETI. (And there is no need to have sense organs that "see" in the radio region. Physics is enough.)

Since perturbation and collision in the asteroid and comet belts are perpetual, the asteroid and comet threat is likewise perpetual, and there is no time when the technology can be retired. Also, SETI itself is a small fraction of the cost of dealing with the asteroid and comet threat.

(Incidentally, it is by no means true that SETI is "very limited, reaching only part of our galaxy." If there were sufficiently powerful transmitters, we could use SETI to explore distant galaxies; because the most likely transmitters are ancient, we can expect them to be powerful. This is one of the strategies of the Megachannel Extraterrestrial Assay [META].)

Is SETI a Fantasy of Physical Scientists?

Mayr has repeatedly suggested that proponents of SETI are almost exclusively physical scientists and that biologists know better. Since the relevant technologies involve the physical sciences, it is reasonable that astronomers, physicists and engineers play a leading role in SETI.

But in 1982, when I put together a petition published in *Science* urging the scientific respectability of SETI, I had no difficulty getting a range of distinguished biologists and biochemists to sign, including David Baltimore, Melvin Calvin, Francis Crick, Manfred Eigen, Thomas Eisner, Stephen Jay Gould, Matthew Meselson, Linus Pauling, David Raup and Edward O. Wilson. In my early speculations on these matters, I was much encouraged by the strong support from my mentor in biology, H. J. Muller, a Nobel laureate in genetics.

The petition proposed that, instead of arguing the issue, we look: "We are unanimous in our conviction that the only significant test of the existence of extraterrestrial intelligence is an experimental one. No a priori arguments on this subject can be compelling or should be used as a substitute for an observational program."

References

L. R. Doyle, ed., *Circumstellar Habitable Zones: Proceedings of the First International Conference*, Travis House Publications, Menlo Park, California, 1996.

Carl Sagan, "The Origin of Life in a Cosmic Context," *Origins of Life*, vol. 5, 1974, pp. 497–505.

Carl Sagan and Ann Druyan, *Shadows of Forgotten Ancestors: A Search for Who We Are*, Random House, New York, 1992.

Carl Sagan et al., "A Search for Life on Earth From the *Galileo* Spacecraft," *Nature*, vol. 365, 1993, pp. 715–721.

Carl Sagan, *Pale Blue Dot: A Vision of the Human Future in Space*, Random House, New York, 1994.

I. S. Shklovskii and Carl Sagan, *Intelligent Life in the Universe*, Holden-Day, San Francisco, 1966.

G. G. Simpson, "The Non-prevalence of Humanoids," *Science*, vol. 143, 1964, pp. 769–775.

To get a handle on the myriad living things that share this planet, humans name and classify every distinct form they identify. They then clump the life-forms into groups. In one system, there are three main branches of life: the eubacteria, the archebacteria and the eukaryota. The eukaryota are divided into four kingdoms— Protista (single-celled or colonial forms), Animalia (our kingdom), Plantae (plants) and Fungi. This last group is highly successful, but its members are not the sort of life-forms that might evolve intelligence. They are, however, relatively close relations of ours; our evolutionary paths diverged a little more than a billion years ago.

Photo: David M. Dennis, Tom Stack & Associates

Typeset and printed in the United Kingdom by
Thanet Press Ltd, Union Crescent, Margate, Kent CT9 1NU
S103s435862i1.5

The Open University

S103 **discovering science**

Supplementary Articles for Block 12

Contents

1.1 DS/P

SUP 74383 0

LIFE'S ROCKY START

Air, water and rock were the only raw materials available on the early earth. The first living entities must have been fabricated from these primitive resources. New experiments suggest that minerals—the basic components of the rocks—could have played starring roles in that dramatic feat.

BY ROBERT M. HAZEN

PHOTOGRAPHS BY ROBERT LEWIS

No one knows how life arose on the desolate young earth, but one thing is certain: life's origin was a chemical event. Once the earth formed 4.5 billion years ago, asteroid impacts periodically shattered and sterilized the planet's surface for another half a billion years. And yet, within a few hundred million years of that hellish age, microscopic life appeared in abundance. Sometime in the interim, the first living entity must have been crafted from air, water and rock.

Of those three raw materials, the atmosphere and oceans have long enjoyed the starring roles in origins-of-life scenarios. But rocks, and the minerals of which they are made, have been called on only as bit players or simply as props. Scientists are now realizing that such limited casting is a mistake. Indeed, a recent flurry of fascinating experiments is revealing that minerals play a crucial part in the basic chemical reactions from which life must have arisen.

The first act of life's origin story must have introduced collections of carbon-based molecules that could make copies of themselves. Achieving even this nascent step in evolution entailed a sequence of chemical transformations, each of which added a level of structure and complexity to a group of organic molecules. The most abundant carbon-based compounds available on the ancient earth were gases with only one atom of carbon per molecule, namely, carbon dioxide, carbon monoxide and methane. But the essential building blocks of living organisms—energy-rich sugars, membrane-forming lipids and complex amino acids—may include more than a dozen carbon atoms per molecule. Many of these molecules, in turn, must bond together to form chainlike polymers and other molecular arrays in order to accomplish life's chemical tasks. Linking small molecules into these complex, extended structures must have been especially difficult in the harsh conditions of the early earth, where intense ultraviolet radiation tended to break down clusters of molecules as quickly as they could form.

Carbon-based molecules needed protection and assistance to enact this drama. It turns out that minerals could have served at least five significant functions, from passive props to active players, in life-inducing chemical reactions. Tiny compartments in mineral structures can shelter simple molecules, while mineral surfaces can provide the scaffolding on which those molecules assemble and grow. Beyond these sheltering and supportive functions, crystal faces of certain minerals can actively select particular molecules resembling those that were destined to become biologically important. The metallic ions in other minerals can jump-start meaningful reactions like those that must have converted simple molecules into self-replicating entities. Most surprising, perhaps, are the recent indications that elements of dissolved minerals can be incorporated into biological molecules. In other words, minerals may not have merely helped biological molecules come together, they might have become part of life itself.

Protection from the Elements

FOR THE BETTER PART of a century, following the 1859 publication of Charles Darwin's *On the Origin of Species*, a parade of scientists speculated on life's chemical origins. Some even had the foresight to mention rocks and minerals in their inventive scenarios. But experimental evidence only sporadically buttressed these speculations.

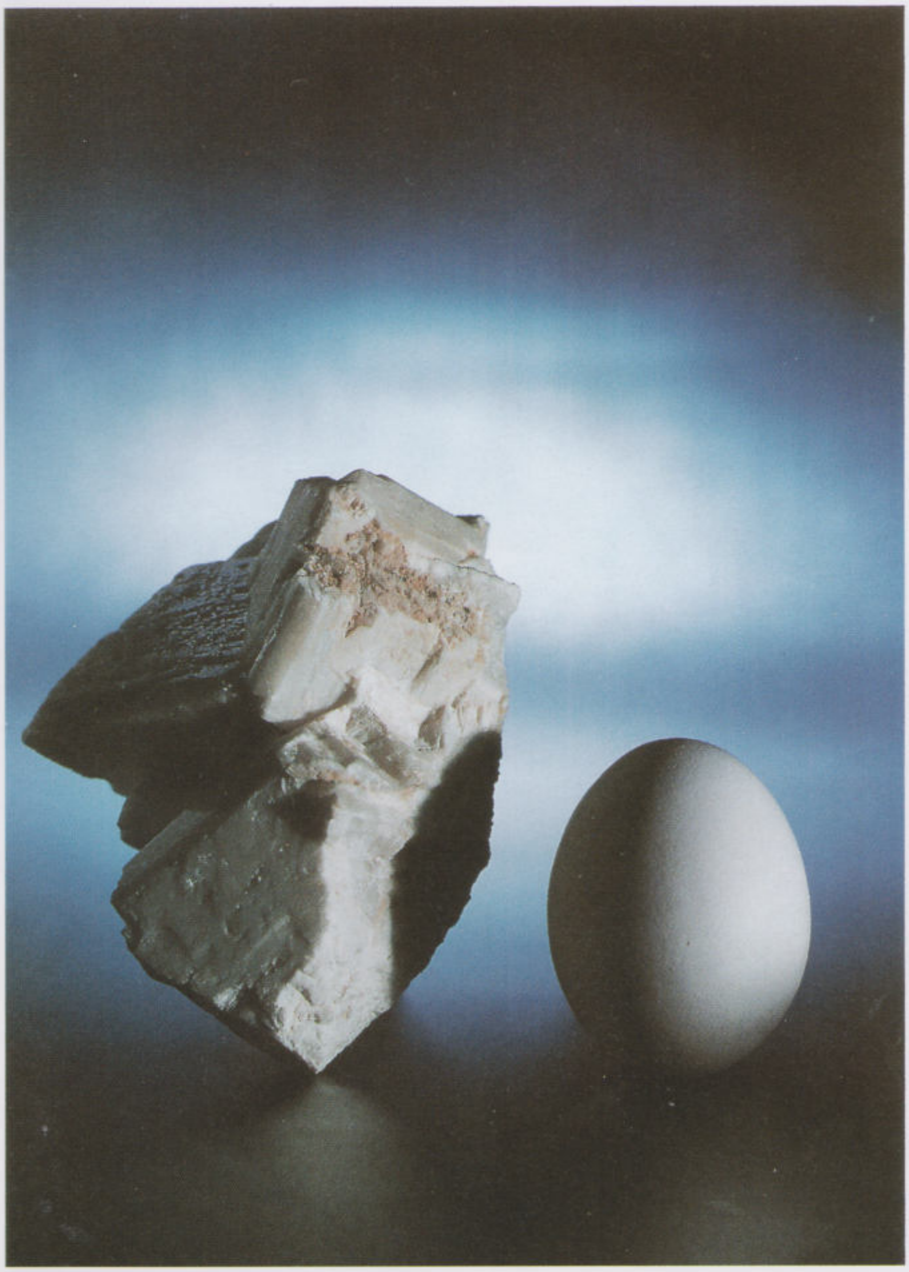
One of the most famous experiments took place at the University of Chicago in 1953. That year chemist Harold C. Urey's precocious graduate student Stanley L. Miller at-

FELDSPAR SPECIMEN COURTESY OF THE AMERICAN MUSEUM OF NATURAL HISTORY, WITH PERMISSION OF THE DEPARTMENT OF EARTH AND PLANETARY SCIENCES

tempted to mimic the earth's primitive oceans and atmosphere in a bottle. Miller enclosed methane, ammonia and other gases thought to be components of the early atmosphere in a glass flask partially filled with water. When he subjected the gas to electric sparks to imitate a prehistoric lightning storm, the clear water turned pink and then brown as it became enriched with amino acids and other essential organic molecules. With this simple yet elegant procedure, Miller transformed origins-of-life research from a speculative philosophical game to an exacting experimental science. The popular press sensationalized the findings by suggesting that synthetic bugs might soon be crawling out of test tubes. The scientific community was more restrained, but many workers sensed that the major obstacle to creating life in the laboratory had been solved.

It did not take long to disabuse researchers of that notion. Miller may have discovered a way to make many of life's building blocks out of the earth's early supply of water and gas, but he had not discovered how or where these simple units would have linked into the complex molecular structures—such as proteins and DNA—that are intrinsic to life.

FELDSPAR: **SHELTERS GROWING CHAINS OF MOLECULES**



To answer that riddle, Miller and other origins scientists began proposing rocks as props. They speculated that organic molecules, floating in seawater, might have splashed into tidal pools along rocky coastlines. These molecules would have become increasingly concentrated through repeated cycles of evaporation, like soup thickening in a heated pot.

In recent years, however, researchers have envisioned that life's ingredients might have accumulated in much smaller containers. Some rocks, like gray volcanic pumice, are laced with air pockets created when gases expanded inside the rock while it was still molten. Many common minerals, such as feldspar, develop microscopic pits during weathering. Each tiny chamber in each rock on the early earth could have housed a separate experiment in molecular self-organization. Given enough time and enough chambers, serendipity might have produced a combination of molecules that would eventually deserve to be called "living."

Underlying much of this speculation was the sense that life was so fragile that it depended on rocks for survival. But in 1977 a startling discovery challenged conventional wisdom about life's fragility and, perhaps, its origins. Until then, most scientists had assumed that life spawned at or near the benign ocean surface as a result of chemistry powered by sunlight. That view began to change when deep-ocean explorers first encountered diverse ecosystems thriving at the superheated mouths of volcanic vents on the seafloor. These extreme environments manage to support elaborate communities of living creatures in isolation from the sun. In these dark realms, much of the energy that organisms need comes not from light but from the earth's internal heat. With this knowledge in mind, a few investigators began to wonder whether organic reactions relevant to the origins of life might occur in the intense heat and pressure of these so-called hydrothermal vents.

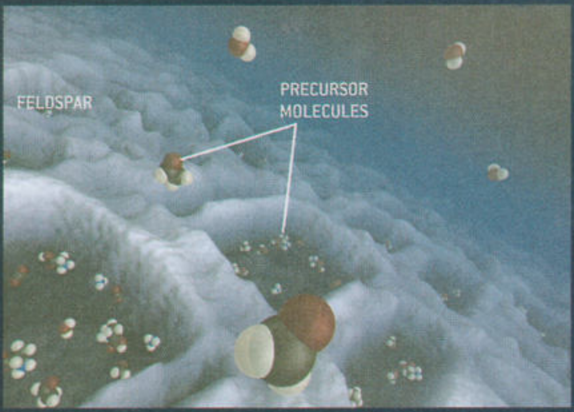
CRYSTAL POWER

NOTHING COULD BE MORE lifeless than a rock, it seems. So how could rocks—or the minerals that constitute them—have assisted the emergence of life? The answer is chemistry. Minerals grow from simple molecules into an ordered structure because of chemical reactions. By the same token, all living organisms—from bacteria to bats—owe their ability to grow and function to the hundreds of chemical reactions that take place inside cells.

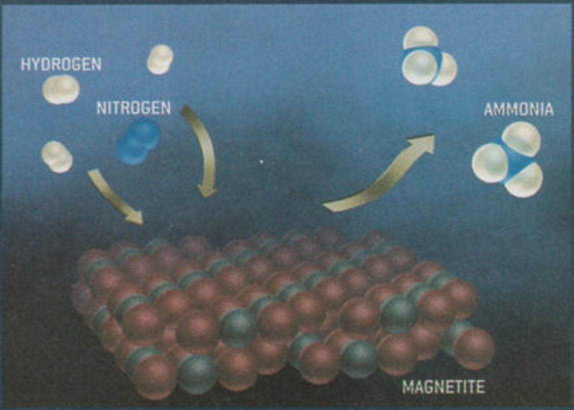
Four billion years ago the earth had no life: chemistry, not biology, altered the planet's surface. In that ancient time minerals—together with the oceans and atmosphere—were the only materials from which the first living entity could have arisen. Chemical reactions, then, must have been the first steps in the origins of life. A sequence of chemical transformations could have reconfigured the simplest components of air, water and rock into primitive collections of carbon-based molecules that could make copies of themselves.

New experiments are revealing that the critical transformations might not have been possible without the help of minerals acting as containers, scaffolds, templates, catalysts and reactants.

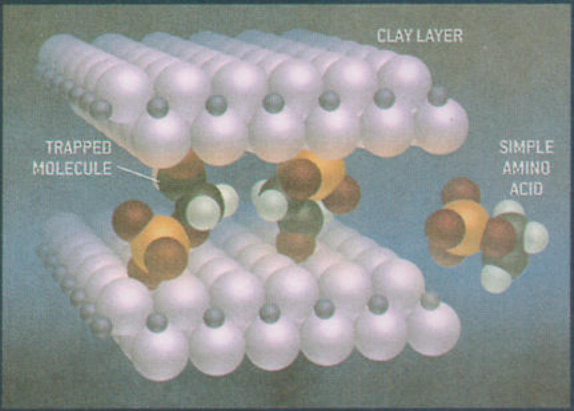
TEMPLATES—The mineral calcite tends to attract left- and right-handed amino acids to different crystal faces. Such a sorting process could explain why life makes use of only the left-handed variety.



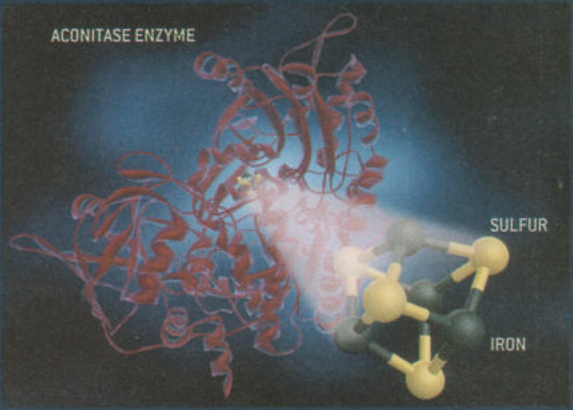
CONTAINERS—Microscopic pits appear in abundance on the weathered surfaces of feldspar and other common minerals. These tiny chambers could have sheltered life's precursor molecules from deadly radiation.



CATALYSTS—Magnetite, an iron oxide mineral, can trigger the recombination of nitrogen and hydrogen gases into ammonia, the essential compound from which living cells acquire nitrogen.



SCAFFOLDS—Layered minerals such as clays can trap stray organic molecules between their rigid sheets of atoms. Held close to one another, simple molecules can react to form more complex compounds.



REACTANTS—Iron and sulfur, the elements that form the active center of certain biological enzymes such as aconitase, can be dissolved from iron sulfide minerals under extreme heat and pressure.

ILLUSTRATIONS BY KENNETH EWARD *BioGrafx*

LAYERED MINERAL: SERVES AS A SCAFFOLD FOR GROWING MOLECULES

Miller and his colleagues have objected to the hydrothermal origins hypothesis in part because amino acids decompose rapidly when they are heated. This objection, it turns out, may be applicable only when key minerals are left out of the equation. The idea that minerals might have sheltered the ingredients of life received a boost from recent experiments conducted at my home base, the Carnegie Institution of Washington's Geophysical Laboratory. As a postdoctoral researcher at Carnegie, my colleague Jay A. Brandes (now at the University of Texas Marine Sciences Institute in Port Aransas) proposed that minerals help delicate amino acids remain intact. In 1998 we conducted an experiment in which the amino acid leucine broke down within a matter of minutes in pressurized water at 200 degrees Celsius—just as Miller and his colleagues predicted. But when Brandes added to the mix an iron sulfide mineral of the type commonly found in and around hydrothermal vents, the amino acid stayed intact for days—plenty of time to react with other critical molecules.

A Rock to Stand On

EVEN IF THE RIGHT raw materials were contained in a protected place—whether it was a tidal pool, a microscopic pit in a mineral surface or somewhere inside the plumbing of a seafloor vent—the individual molecules would still be suspended in water. These stray molecules needed a support structure—some kind of scaffolding—where they could cling and react with one another.

Simple molecules could have used RIGID MINERAL SURFACES as the scaffolding on which they reassembled into more complex structures.

One easy way to assemble molecules from a dilute solution is to concentrate them on a flat surface. Errant molecules might have been drawn to the calm surface of a tidal pool or perhaps to a primitive "oil slick" of compounds trapped at the water's surface. But such environments would have posed a potentially fatal hazard to delicate molecules. Harsh lightning storms and ultraviolet radiation accosted the young earth in doses many times greater than they do today. Such conditions would have quickly broken the bonds of complex chains of molecules.

Origins scientists with a penchant for geology have long recognized that minerals might provide attractive alternative surfaces where important molecules could assemble. Like the container idea, this notion was born half a century ago. At that time, a few scientists had begun to suspect that clays have special abilities to attract organic molecules [*see box on*

page 80]. These ubiquitous minerals feel slick when wet because their atoms form flat, smooth layers. The surfaces of these layers frequently carry an electric charge, which might be able to attract organic molecules and hold them in place. Experiments later confirmed these speculations. In the late 1970s an Israeli research group demonstrated that amino acids can concentrate on clay surfaces and then link up into short chains that resemble biological proteins. These chemical reactions occurred when the investigators evaporated a water-based solution containing amino acids from a vessel containing clays—a situation not unlike the evaporation of a shallow pond or tidal pool with a muddy bottom.

More recently, separate research teams led by James P. Ferris of the Rensselaer Polytechnic Institute and by Gustaf Arrhenius of the Scripps Institution of Oceanography demonstrated that clays and other layered minerals can attract and assemble a variety of organic molecules. In a tour de force series of experiments during the past decade, the team at Rensselaer found that clays can act as scaffolds for the building blocks of RNA, the molecule in living organisms that translates genetic instructions into proteins.

Once organic molecules had attached themselves to a mineral scaffold, various types of complex molecules could have been forged. But only a chosen few were eventually incorporated into living cells. That means that some kind of template must have selected the primitive molecules that would become biologically important. Recent experiments show, once again, that minerals may have played a central role in this task.

Preferential Treatment

PERHAPS THE MOST mysterious episode of selection left all living organisms with a strange predominance of one type of amino acid. Like many organic molecules, amino acids come in two forms. Each version comprises the same types of atoms, but the two molecules are constructed as mirror images of each other. The phenomenon is called chirality, but for simplicity's sake scientists refer to the two versions as "left-handed" (or "L") and "right-handed" (or "D"). Organic synthesis experiments like Miller's invariably produce 50–50 mixtures of L and D molecules, but the excess of left-handed amino acids in living organisms is nearly 100 percent.

Researchers have proposed a dozen theories—from the mundane to the exotic—to account for this bizarre occurrence. Some astrophysicists have argued that the earth might have formed with an excess of L amino acids—a consequence of processes that took place in the cloud of dust and gas that became the solar system. The main problem with this theory is that in most situations such processes yield only the slightest excess—less than 1 percent—of L or D molecules.

Alternatively, the world might have started with a 50–50 mixture of L and D amino acids, and then some important feature of the physical environment selected one version over the other. To me, the most obvious candidates for this specialized physical environment are crystal faces whose surface structures are mirror images of each other [*see box on page 80*]. Last spring I narrowed in on calcite, the common mineral that forms limestone and marble, in part because it often displays magnificent pairs of mirror-image faces. The chemical structure of calcite in many mollusk shells bonds strongly to amino acids. Knowing this, I began to suspect that calcite surfaces may feature chemical bonding sites that are ideally suited to only one type of amino acid or the other. With the help of my Carnegie colleague Timothy Filley (now at Purdue University) and Glenn Goodfriend of George Washington University, I ran more than 100 tests of this hypothesis.

Our experiments were simple in concept, although they required meticulous clean-room procedures to avoid contamination by the amino acids that exist everywhere in the environment. We immersed a well-formed, fist-size crystal of calcite into a 50–50 solution of aspartic acid, a common amino acid. After 24 hours we removed the crystal from this solution, washed it in water and carefully collected all the molecules that had adhered to specific crystal faces. In one experiment after another we observed that calcite's "left-handed" faces selected L-amino acids, and vice versa, with excesses approaching 40 percent in some cases.

Curiously, calcite faces with finely terraced surfaces displayed the greatest selectivity. This outcome led us to speculate that these terraced edges might force the L and D amino acids to line up in neat rows on their respective faces. Under the right environmental conditions, these organized rows of amino acids might chemically join to form proteinlike molecules—some made entirely of L

Crystal faces of certain minerals could have ACTIVELY SELECTED and concentrated molecules that were destined to become biologically important.

THE AUTHOR

ROBERT M. HAZEN has explored the behavior of minerals under high pressure at the Carnegie Institution of Washington's Geophysical Laboratory since 1976. In the past five years he has designed many of his mineral experiments to mimic the high-pressure environments of deep-sea hydrothermal vents. Rocks and minerals first piqued Hazen's curiosity as a child growing up in northern New Jersey, a region known for its unusual ore deposits. After receiving a doctorate in earth sciences at Harvard University in 1975 and spending a year at the University of Cambridge, he joined the staff at Carnegie. In 1990 Hazen took on a second position, as professor of earth science at George Mason University. He is also a part-time professional trumpeter and the author of numerous articles and books on science, education, history and music.

CALCITE: SELECTS FROM MIRROR-IMAGE MOLECULES

CALCITE SPECIMEN COURTESY OF LAWRENCE H. CONKLIN

amino acids, others entirely of D. If protein formation can indeed occur, this result becomes even more exciting, because recent experiments by other investigators indicate that some proteins can self-replicate. In the earth's early history, perhaps a self-replicating protein formed on the face of a calcite crystal.

Left- and right-handed crystal faces occur in roughly equal numbers, so chiral selection of L amino acids probably did not happen everywhere in the world at once. Our results and predictions instead suggest that the first successful set of self-replicating molecules—the precursor to all the varied life-forms on the earth today—arose at a specific time and place. It was purely chance that the successful molecule developed on a crystal face that preferentially selected left-handed amino acids over their right-handed counterparts.

Minerals undoubtedly could have acted as containers, scaffolds and templates that helped to select and organize the molecular menagerie of the primitive earth. But many of us in origins research suspect that minerals played much more active roles, catalyzing key synthesis steps that boosted the earth's early inventory of complex biological molecules.

Getting a Jump on the Action

EXPERIMENTS LED by Carnegie researcher Brandes in 1997 illustrate this idea. Biological reactions require nitrogen in the form of ammonia, but the only common nitrogen compound thought to have been available on the primitive earth is nitrogen gas. Perhaps, Brandes thought, the environment at hydrothermal vents mimics an industrial process in which ammonia is synthesized by passing nitrogen and hydrogen over a hot metallic surface. Sure enough, when we subjected hydrogen, nitrogen and the iron oxide mineral magnetite to the pressures and temperatures characteristic of a seafloor vent, the mineral catalyzed the synthesis of ammonia [*see box on page 80*].

The idea that minerals may have triggered life's first crucial steps has emerged most forcefully from the landmark theory of chemist Günter Wächtershäuser, a German patent lawyer with a deep interest in life's origins. In 1988 Wächtershäuser advanced a sweeping theory of organic evolution in which minerals—mostly iron and nickel sulfides that abound at deep-sea hydrothermal vents—could have served as the template, the catalyst and the energy source that drove the formation of biological molecules. Indeed, he has argued that primitive living entities were molecular coatings that adhered to the positively charged surfaces of pyrite, a mineral composed of iron and sulfur. These entities, he further suggests, obtained energy from the chemical reactions that produce pyrite. This hypothesis makes sense in part because some metabolic enzymes—the molecules that help living cells process energy—have at their core a cluster of metal and sulfur atoms.

For much of the past three years, Wächtershäuser's provocative theory has influenced our experiments at Carnegie. Our team, including geochemist George Cody and petrologist Hatten S. Yoder, has focused on the possibility that metabolism can proceed without enzymes in the presence of minerals—especially oxides and sulfides. Our simple strategy,

MORE TO EXPLORE

Beginnings of Cellular Life. Harold J. Morowitz. Yale University Press, 1992.

Origins of Life: The Central Concepts. David W. Deamer and Gail R. Fleischaker. Jones and Bartlett, 1994.

Emergence: From Chaos to Order. John H. Holland. Helix Books, 1998.

Biogenesis: Theories of Life's Origin. Noam Lahav. Oxford University Press, 1999.

much in the spirit of Miller's famous experiment, has been to subject ingredients known to be available on the young earth—water, carbon dioxide and minerals—to a controlled environment. In our case, we try to replicate the bone-crushing pressures and scalding temperatures typical of a deep-sea hydrothermal vent. Most of our experiments test the interactions among ingredients enclosed in welded gold capsules, which are roughly the size of a daily vitamin pill. We place as many as six capsules into Yoder's "bomb"—a massive steel pressure chamber that squeezes the tiny capsules to pressures approaching 2,000 atmospheres and heats them to about 250 degrees C.

One of our primary goals in these organic-synthesis experiments—and one of life's fundamental chemical reactions—is carbon fixation, the process of producing molecules with an increasing number of carbon atoms in their chemical structure. Such reactions follow two different paths depending on the mineral we use. We find that many common minerals, including most oxides and sulfides of iron, copper and zinc, promote carbon addition by a routine industrial process known as Fischer-Tropsch (F-T) synthesis.

Minerals could have jump-started CRITICAL CHEMICAL REACTIONS that boosted the earth's early inventory of complex biological molecules.

This process can build chainlike organic molecules from carbon monoxide and hydrogen. First, carbon monoxide and hydrogen react to form methane, which has one carbon atom. Adding more carbon monoxide and hydrogen to the methane produces ethane, a two-carbon molecule, and then the reaction repeats itself, adding a carbon atom each time. In the chemical industry, researchers have harnessed this reaction to manufacture molecules with virtually any desired number of carbon atoms. Our first organic-synthesis experiments in 1996, and much more extensive research by Thomas McCollom of the Woods Hole Oceanographic Institution, demonstrate that F-T reactions can build molecules with 30 or more carbon atoms under some hydrothermal-vent conditions in less than a day. If this process manufactures large organic molecules from simple inorganic chemicals throughout the earth's hydrothermal zones today, then it very likely did so in the planet's prebiological past.

When we conduct experiments using nickel or cobalt sulfides, we see that carbon addition occurs primarily by carbonylation—the insertion of a carbon and oxygen molecule, or carbonyl group. Carbonyl groups readily attach themselves to nickel or cobalt atoms, but not so strongly that they cannot link to other molecules and jump ship to form larger molecules. In one series of experiments, we observed the lengthening of the nine-carbon molecule nonyl thiol to form 10-carbon decanoic acid, a compound similar to the acids that drive metabolic reactions in living cells. What is more, all the reactants in this experiment—a thiol, carbon monoxide and water—are readily available near sulfide-rich hydrothermal vents. By repeating these simple kinds of reactions—adding a carbonyl group here or a hydroxide group there—we can synthesize a rich variety of complex organic molecules.

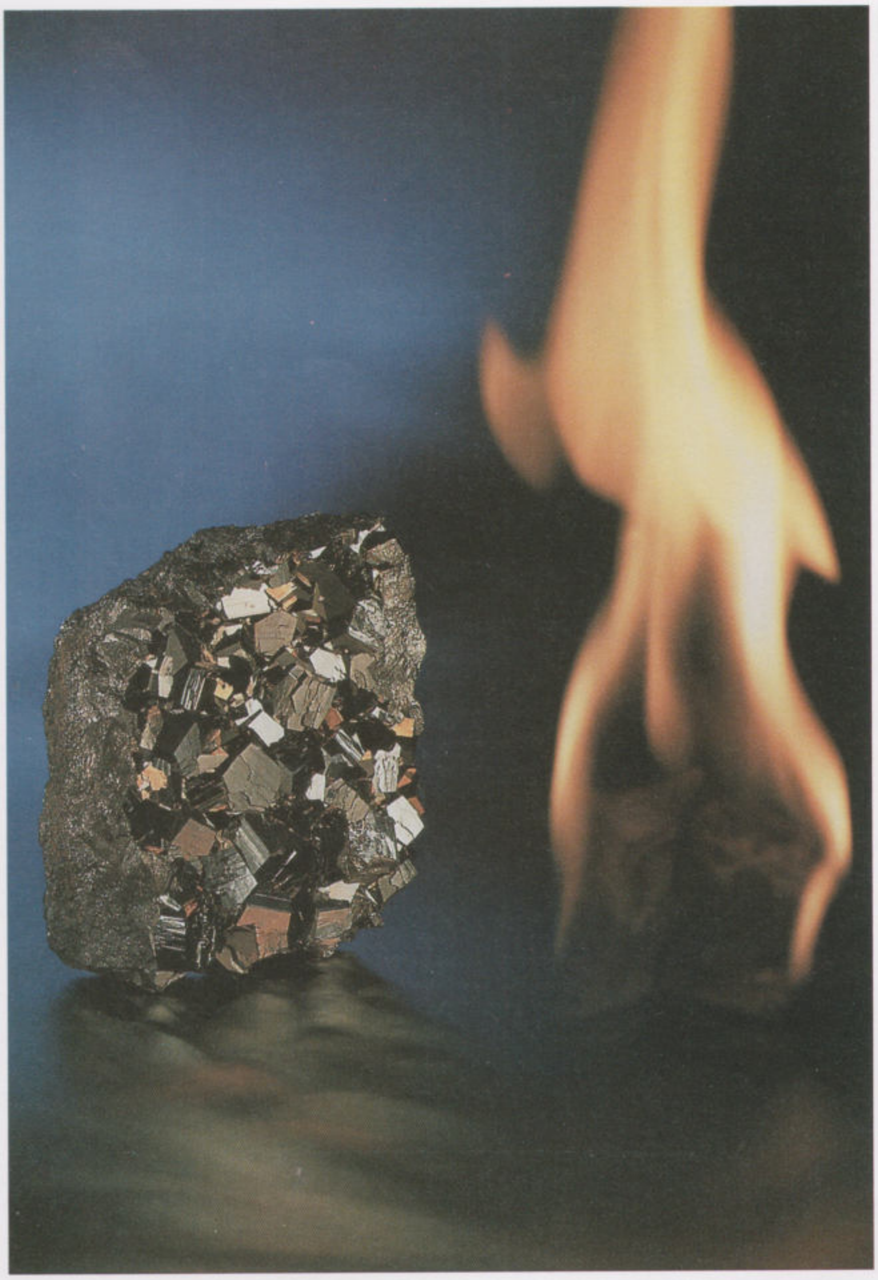
Our 1,500 hydrothermal organic synthesis experiments at Carnegie have done more than supplement the catalogue of interesting molecules that must have been produced on the early earth. These efforts reveal another, more complex behavior of minerals that may have significant consequences for the chemistry of life. Most previous origins-of-life studies have treated minerals as solid and unchanging—stable platforms where organic molecules could assemble. But we are finding that in the presence of hot water at high pressure, minerals start to dissolve. In the process, the liberated atoms and molecules from the minerals can become crucial reactants in the primordial soup.

The Heart of the Matter

OUR FIRST DISCOVERY of minerals as reactants was an unexpected result of our recent catalysis experiments led by Cody. As expected, carbonylation reactions produced 10-carbon decanoic acid from a mixture of simple molecules inside our gold capsules. But significant

MAGNETITE: CATALYZES BIOCHEMICAL REACTIONS

MAGNETITE SPECIMEN COURTESY OF LAWRENCE H. CONKLIN



PYRITE: **FUELS BIOCHEMICAL REACTIONS**

quantities of elemental sulfur, organic sulfides, methyl thiol and other sulfur compounds appeared as well. The sulfur in all these products must have been liberated from the iron sulfide mineral.

Even more striking was the liberation of iron, which brilliantly colored the water-based solutions inside the capsules. As the mineral dissolved, the iron formed bright red and orange organometallic complexes in which iron atoms are surrounded by various organic molecules. We are now investigating the extent to which these potentially reactive complexes might act as enzymes that promote the synthesis of molecular structures.

The role of minerals as essential chemical ingredients of life is not entirely unexpected. Hydrothermal fluids are well known to dissolve and concentrate mineral matter. At deep-sea vents, spectacular pillars of sulfide grow dozens of feet tall as plumes of hot, mineral-laden water rise from below the seafloor, contact the frigid water of the deep ocean and deposit new layers of minerals on the growing pillar. But the role of these dissolved minerals has not yet figured significantly in origins scenarios. Whatever their behavior, dissolved minerals seem to make the story of life's emergence much more interesting.

When we look beyond the specifics of prebiological chemistry, it is clear that the origin of life was far too complex to imagine as a single event. Rather we must work from the assumption that it was a gradual sequence of more modest events, each of which added a degree of order and complexity to the world of prebiological molecules. The first step must have been the synthesis of the basic building blocks. Half a century of research reveals that the molecules of life were manufactured in abundance—in the nebula that formed our solar system, at the ocean's surface, and near hydrothermal vents. The ancient earth suffered an embarrassment of riches—a far greater diversity of molecules than life could possibly employ.

Minerals helped to impose order on this chaos. First by confining and concentrating molecules, then by selecting and arranging those molecules, minerals may have jump-started the first self-replicating molecular systems. Such a system would not have constituted life as we know it, but it could have, for the first time, displayed a key property of life. In this scenario, a self-replicating molecular system began to use up the resources of its environment. As mutations led to slightly different variants, competition for limited resources initiated and drove the process of molecular natural selection. Self-replicating molecular systems began to evolve, inevitably becoming more efficient and more complex.

A long-term objective for our work at the Carnegie Institution is to demonstrate simple chemical steps that could lead to a self-replicating system—perhaps one related to the metabolic cycles common to all living cells. Scientists are far from creating life in the laboratory, and it may never be possible to prove exactly what chemical transformations gave rise to life on earth. What we can say for sure is that minerals played a much more complex and integral part in the origin of life than most scientists ever suspected. By being willing to cast minerals in starring roles in experiments that address life's beginnings, researchers may come closer to answering one of science's oldest questions.

Lifting Titan's Veil: Exploring the giant moon of Saturn

Extract 1

Discovering Titan

The landscape seems alien. The few clouds that burned a garish red as the Sun set have flitted away and the sky is clear. Strange and unfamiliar life-forms, deprived of water, struggle to survive in the harsh conditions. This is no extraterrestrial scene though, but Tucson, Arizona. Arizona's commendably dark and clear skies are a magnetic attraction for astronomers.

High above, Jupiter gleams brilliantly. Through even a small telescope, an entourage of four moons circling this giant planet and its cloud bands alternating light and dark are obvious. A short distance to the east there is another planet, not as bright as Jupiter. It's Saturn. Through the telescope it is an altogether different object, with its rings tilted tastefully – as though a jeweller had set it there. A little to one side of the rings is a dim, unprepossessing dot, looking a little reddish maybe. The dot is the focus of our attention – Saturn's moon, Titan, a world as intriguing as any in the solar system.

On the 15th of October 1997, another Titan roared into the sky. To be precise, it was a Titan IVB/Centaur launch vehicle. Just before 5 o'clock in the morning local time the appropriately named rocket blasted off from Cape Canaveral Air Force Station, Florida, bearing a 5.8-tonne spacecraft bound for Saturn and Titan. It was the start of a seven-year journey for the *Cassini-Huygens* mission and of a tantalising seven-year wait for the anxious scientists on the ground. *Cassini* was destined to enter orbit around Saturn on the 1st of July 2004. Seven months later, if all goes according to plan, the *Huygens* probe will detach itself, cruise towards Titan for about three weeks, then parachute down onto Titan's surface. Instruments on board the *Cassini* orbiter will gather data about Saturn and its moons, especially Titan, over a four-year period. Then Titan will become the most distant world by far to have a human artefact land upon it. The enormous effort dedicated to achieving this feat is a testament to the growth in our fascination with Titan as work of unique significance in the quest to understand our own planet.

Figure 1.1 The launch of the ***Cassini-Huygens*** mission on the 15th of October 1997 at 4.43 a.m. EDT, from Cape Canaveral Air Station in Florida. The launch vehicle was a Titan IVB/Centaur. NASA image.

Extract 2

The methane fingerprint

With rare exceptions, astronomers can never hope to get their hands on material from the objects they study. But even though samples of matter are out of the question, energy samples arrive unsolicited. Lights and its invisible relatives – ultraviolet and infrared radiation and radio waves, for example – are there for the taking with the right telescopes and instruments. Fortunately, many of the substances in stars, planets, or indeed anything that shines, indelibly impress their unique fingerprints on the energy they radiate away.

Extracting the tell-take fingerprints from a beam of light is achieved by spectroscopy – a process rather like disentangling a vast skein of multicoloured yarn of various lengths. Each colour corresponds to a different wavelength and the length of thread represents the intensity of the light in that individual colour. The signature of a particular chemical is a unique combination of threads with certain lengths and colours.

For scientific analysis, a spectrum has to be transformed into a trace, displaying the ups and downs of light intensity with wavelength. Common features in such spectrum traces have their own descriptive expressions in the scientific jargon. A series of closely spaced narrow dips, which may appear to be merged together to make a single broad dip, is an 'absorption band', for example. The fingerprints of molecules typically take the guise of absorption bands.

In the winter of 1943-44, the Dutch-American astronomer Gerard P Kuiper used the 82-inch (2.08-m) telescope at the McDonald Observatory in Texas to record the spectra of the ten largest moons in the solar system in both visible light and the near infrared. This exercise was challenging at that time and had not been attempted before. The telescope was still relatively new, having been completed in 1938.

Kuiper's spectrum of Titan immediately stood out from the rest. Uniquely, it contained absorption bands identified with methane gas. Titan's orange hue was also apparent in the data, confirming what many an observer had noted by eye. Though a link with the atmosphere was an obvious connection to make, Kuiper could not know at that time whether the surface of Titan or the atmosphere itself was responsible for Titan's distinctive colour: Kuiper published his results in a paper under the title *Titan: A satellite with an atmosphere*. While James Jeans had shown decades earlier that physics and chemistry allowed Titan to have an atmosphere, the proof of its existence was still for many a startling revelation.

Unlike Comas Solà's obscure and unrepeated observation, Kuiper's evidence was clear and indisputable: Titan was no run-of-the-mill satellite. Yet the strong signature of methane was far from the last word on Titan's atmosphere. The bigger story would have to wait until 1980 and the arrival of the *Voyager 1* spacecraft. Now we know that methane, which trumpets its presence with a strong

spectral signature, accounts for a mere few per cent of Titan's atmosphere at most. Its publicity-shy but more abundant partner turned out to be nitrogen. As in Earth's atmosphere, the nitrogen atoms pair up to form molecules, N_2. Together, nitrogen and methane weight down on Titan with a pressure one and half times greater than atmospheric pressure on Earth.

Voyager 1 reached Saturn in November 1980 after a journey lasting just over 13 years. Its encounter with Titan was brief but intimate. The craft closed in to a mere 4394 km from the surface at nearest approach, its camera and full arsenal of instruments trained on the mysterious moon. In the 1970s, before *Voyager 1*'s arrival, astronomers began to suspect that Titan's atmosphere contains clouds and haze. Yet they held out the hope of gaps – a glimpse perhaps of the surface below. But it was not to be. *Voyager*'s camera, sensitive to visible light, returned images of a moon comprehensively swathed in a global blanket of orange smog. It was a disappointment but not so great a surprise. Ultraviolet light from the Sun breaks up the molecules of methane and nitrogen in Titan's upper atmosphere, releasing the ingredients to cook up a soup of complex chemicals. Some of the new substances created from the dismembered fragments are probably polymers – large, chain-like molecules. According to *Voyager 1* data, the dark particles suspended high above Titan's surface are about 0.2 to 1.0 micron across. No-one knows the details of their chemistry for sure. Whatever their nature, they are guilty of concealment and of provoking intrigue on another world a billion miles away.

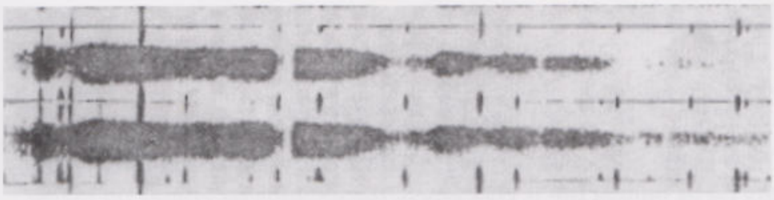


Figure 1.7 Two of Gerald Kuiper's photographic spectra of Titan, taken in the winter of 1943–44 with the 82-inch (2.08 m) telescope at the McDonald Observatory in Texas. They were published in the *Astrophysical Journal* in 1944 (vol. **100**, p. 378). Wavelength increases from left to right. They are negatives, so the methane absorption bands, which disclosed the presence of an atmosphere on Titan, appear as light gaps in the broad, dark, horizontal strips. The short vertical lines are wavelength reference marks.

A singular satellite

Titan's atmosphere is distinctive, fascinating and unique in the solar system, imbuing its owner with the qualities of a true planetary world according to all our preconceptions of what moons and planets should be like. Reinforcing this notion of planetary status, the composition of Titan's nitrogen-rich atmosphere is beguilingly similar to Earth's and unlike every other substantial atmosphere in the solar system. Not only that, the pressure and density of the atmospheres surrounding Earth and Titan are similar. No other world's atmosphere matches Earth's so closely in this respect. But there the similarities between Earth and Titan end. The trace gases in the two atmospheres are very different, largely because of the low temperature on Titan.

As a moon of Saturn, Titan is firmly in the 'outer' solar system. This means that it is cold – very cold. Its surface is 180°C below zero – or 94 K on the absolute temperature scale preferred by scientists. It means also that Titan is made of a mix of ingredients different from those familiar on Earth and the other terrestrial planets, which are made almost entirely of rock and metal. Titan, like most of the other bodies in the outer solar system, is composed mainly of ice. Carbon dioxide and water vapour, minor but significant gases in Earth's atmosphere, are frozen solid on Titan. Instead, Titan's atmosphere boasts a cocktail of carbon-based chemicals, including ethane, acetylene and carbon monoxide.

In some respects Titan is like the planet Venus. Both have atmospheres that are thick and opaque to sunlight and both rotate slowly. This combination appears to lead to fast winds in the upper atmosphere, which zip around from east to west. In another sense, Titan is faintly reminiscent of Mars, in that the tilt of its equator results in pronounced seasons and the movement of atmospheric gas from the summer hemisphere to the winter one. On Mars, where the atmosphere is thin, the effect is huge: carbon dioxide frost evaporates in the summer hemisphere and snows out on the winter hemisphere. Atmospheric pressure changes by around 30%. On Titan, with its much thicker atmosphere, the effect is much more subtle. The haze high in Titan's atmosphere seems to be driven from the summer hemisphere to the winter one – changing Titan's brightness quite dramatically as it moves.

Extract 3

Titan's puzzling atmosphere

How many conversations begin with a discussion on the weather? As far as human life is concerned, the tumultuous goings on in the atmosphere of our planet Earth are of immense consequence, from day to day, throughout the year and over the centuries. And we devote considerable resources to recording and predicting the ever-changing behaviour of our blanket of gas. We worry about the causes of global warming and thinning of the ozone layer. Worlds with atmospheres have an extra dimension of complexity over those that do not, an extra layer to be analysed, characterised and understood.

And so it is with Titan. Its surface may be hidden but between us and that secret landscape lies an amazing layer of gas. Can there really be a moon with such an atmosphere? It's still a novel phenomenon. What is this stuff? What is it doing there?

In Extract 1 we described how Titan's atmosphere was properly discovered in 1944 (if you don't count Comas Solà's irreproducible 1908 report). In this extract, we'll go into more detail about what it's made of, how we think it was created (admitting right away that no-one knows for sure), and how it has evolved. Most intriguing of all is what Titan's atmosphere might be able to tell us about our own atmosphere and even, perhaps, the origin of life on Earth.

The mystery gas

Back in the mid-1970s, there were competing ideas about the unknown composition of Titan's atmosphere. After Kuiper's discovery of methane, the next substance to be identified in the atmosphere was hydrogen, found by Larry Trafton. In some respects this was no surprise, since hydrogen dominates the atmospheres of all the outer planets. On the other hand, hydrogen is a very light molecule and, as James Jeans noted in 1925, it shouldn't hang around on a body with gravity as weak as Titan's. Its existence did present a puzzle.

Careful study of spectra suggested that there was something else there besides the hydrogen and methane. Ammonia was discussed at some length as a possible candidate and, without knowing how low the surface temperature was, there was no way of ruling it out. Nitrogen was considered as an outside contender, although it was favoured by Tucson scientist Donald Hunten at the 1973 workshop on Titan.

Even in those days, researchers knew that methane would undergo chemical changes as a result of being bathed in ultraviolet light, a process called 'photochemistry'. A pioneering experiment had been undertaken by Harold Urey and Stanley Miller in the 1950s: they took a mixture of gases and water vapour, kept it warm and passed an electric spark through it for a few days. The liquid in the flask turned a little pink at first and eventually became a tarry sludge. When they analysed its contents they found many important organic molecules. Later experiments found that similar results could be obtained by exposing gas mixtures to ultraviolet light. A more recent version of this experiment, using a laser as the energy source, is illustrated in Figure 3.1.

Exactly what the products of photochemistry in Titan's atmosphere would be no-one knew. 'Asphalt' was the catch-all term used by the 1973 workshop participants. Carl Sagan reported early results of experiments of his own, following Urey and Miller, in which electric sparks were passed through methane gas mixtures. This produced a reddish sludge, for which Sagan coined the term 'tholin' inspired by the Greek word for mud. Sagan argued that something of the kind could account for the orange colour of Titan. As a haze, it might be responsible for the unusual manner in which Titan reflected light and for the apparently warm stratosphere.

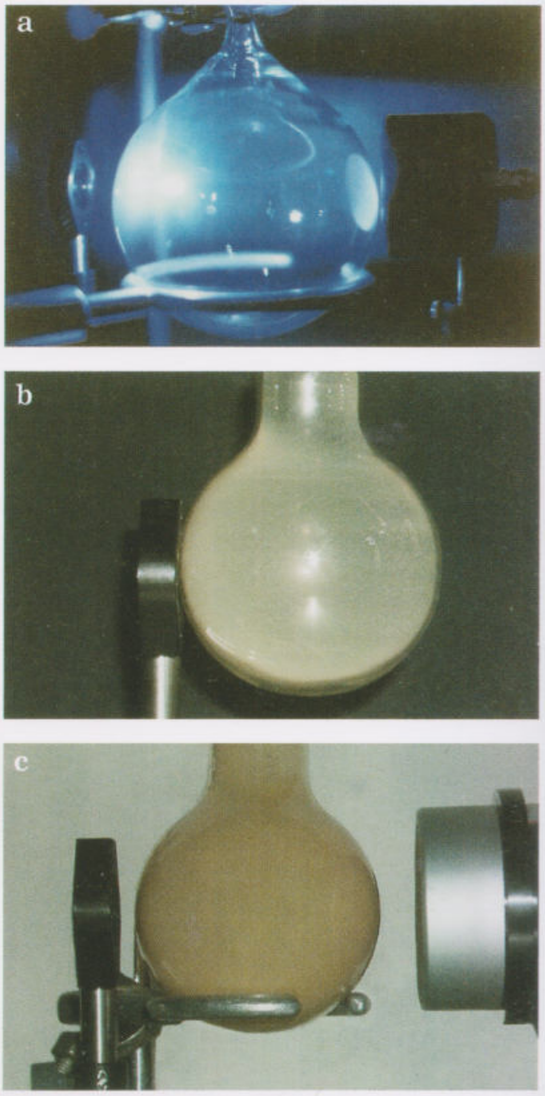


Figure 3.1 Simulating lightning on Titan. In the experiment pictured here, which was carried out in 2001, a focused (invisible) laser beam was used to make a white-hot plasma in a flask of methane and nitrogen (a). After a short time, the flask became fogged with a white deposit (b). The deposit became brownish as it thickened (c). The brown deposit of organic chemicals resembles Titan's haze. Laser plasma experiments are more convenient than the electrical discharge and ultraviolet illumination experiments conducted by Miller, Sagan and others because the deposit is generated in only a few hours rather than days. Images courtesy of Rafael Navarro-González, Laboratory of Plasma Chemistry and Planetary Studies, National University of Mexico.

The revelations of Voyager

When *Voyager 1* arrived in the saturnian system in 1980, it was carrying an instrument that immediately revealed the identity of the mystery constituent in Titan's atmosphere. The ultraviolet spectrometer unambiguously detected the faint glow of nitrogen. There was the answer: The combination of *Voyager*'s radio occultation experiment, ultraviolet data and infrared spectra indicated that the atmosphere was between 85% and 95% nitrogen, with most of the rest methane. There was about 2% methane in the stratosphere, although near the cold surface, the amount might be nearer 6% or 8%. Argon remained a wild card and then there was one or two tenths of a per cent hydrogen.

But the exact mix of gases remains uncertain to this day. Is there 5% of argon, or 0.5%? It sounds like a very academic question but it bears directly on such disparate concerns as the origin of the atmosphere and the design of the heat shield of the *Huygens* probe. If the argon abundance is at the high end of the range, then the methane abundance near the surface also has to be higher, to account for the average molecular weight of the gas mix there. We know the average molecular weight of Titan's atmosphere is close to that of pure nitrogen – 28 – but argon is a heavy gas with a molecular weight of 40. If argon is plentiful, the average molecular weight is being dragged down by extra methane: methane's light molecules weigh in at only 16. Methane in Titan's atmosphere is the analogue of humidity in Earth's, so the amount is crucially important for the stability of the climate and for the meteorological processes such as rainfall and clouds.

It took over a decade for the infrared data to be fully understood. The myriad of organic compounds took a lot of unravelling. Many effects had to be folded into the analysis – instrumental corrections, the temperature profile between the surface and the stratosphere, the angle of view. Such a detailed and painstaking investigation is usually delegated to a Ph.D. student and this one was no exception. The job fell to Athena Coustenis. By fitting the observed spectra, such as the one in Figure 3.2, to the predictions of a detailed simulation, she was able to calculate the abundance of about a dozen different kinds of molecules present in the atmosphere. Furthermore, she established that the abundances of some of them varied with latitude: ethylene and the more exotic methylacetylene and diacetylene, together with various nitriles, were around ten times more common at high northern latitudes than elsewhere. This observation could be put down to a seasonal phenomenon. The *Voyager* encounter took place during northern spring on Titan and the high northern latitudes had just emerged from a long polar night. They had been in darkness for years. The chemistry here would have been very different without light to drive photochemical reactions and without any warmth from the Sun. At the lower temperatures here, some chemical reactions would be slower than at other parts of the globe and some compounds would condense. In fact, exactly the same situation occurs on Earth. During the polar night, the atmospheric chemistry subtly alters and a rare but important gas becomes severely depleted: the now notorious ozone hole opens up.

The latitudes at which these chemical perturbations occur correspond closely with the dark polar 'hood' apparent in *Voyager*'s blue-filtered images of Titan. Dynamically this region may be a circumpolar vortex, a sort of Sargasso Sea in the sky, where the chemistry is isolated from the rest of the atmosphere by winds slavishly following lines of latitude. This is the kind of circulation pattern that occurs around the Antarctic on Earth, where the more severe ozone hole builds up. Perhaps understanding Earth's polar photochemistry will help us understand Titan's, if not vice versa.

Coustenis also studied infrared spectra from *Voyager 2*. Even though this spacecraft approached Titan no closer than 600 000 km, its spectra could be compared with those from *Voyager 1* obtained nine months earlier. Nine months out of Titan's 30-year seasonal cycle is the equivalent of only one week on Earth. Even so, there were significant variations in the abundance of methylacetylene and diacetylene at high latitudes, perhaps the effects of spring sunlight on the polar hood.

As a result of their analyses, Coustenis and her colleagues Daniel Gautier, Bruno Bézard and Emmanuel Lellouch at the Observatoire de Paris in Meudon, were able to say something about the distribution with height of several of the gases. Since most of them are produced at high altitudes and are removed by condensation lower down, their abundances decrease at lower altitudes. Such details are not of broad interest, perhaps, but they are vital when it comes to distinguishing between different chemical scenarios.

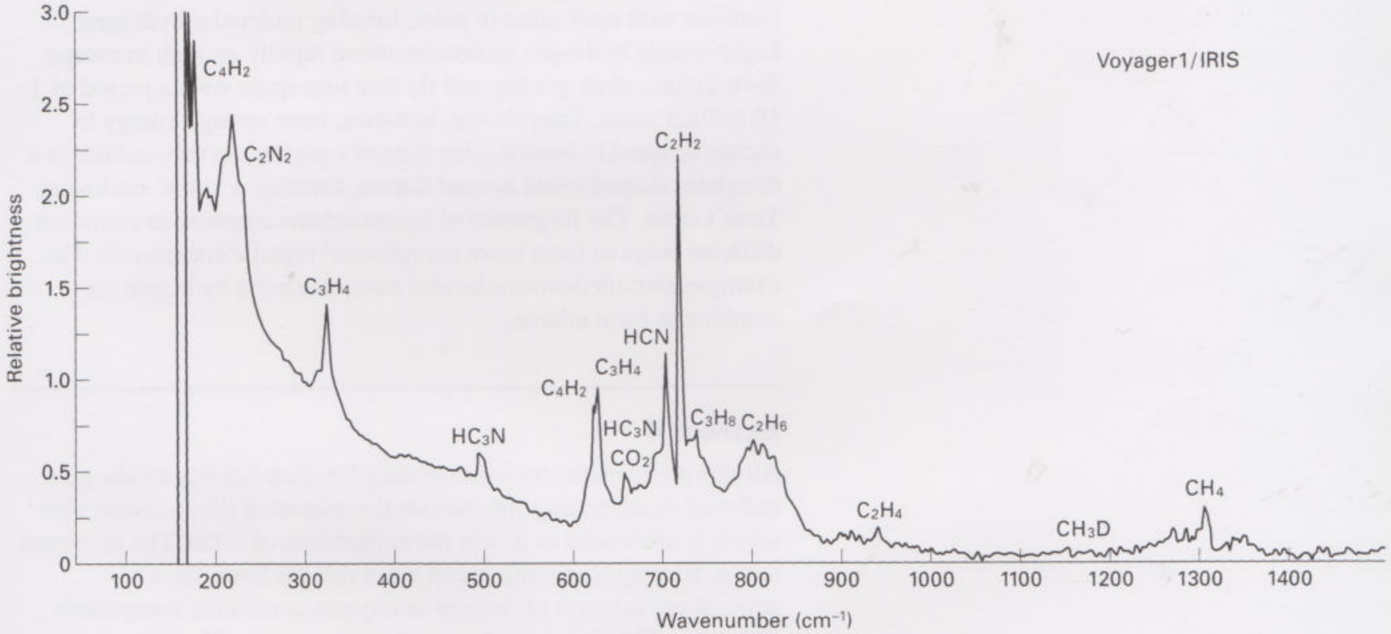
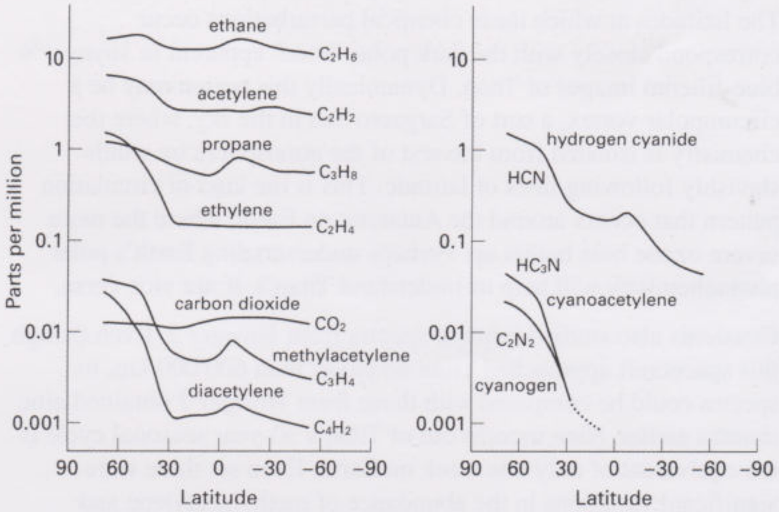


Figure 3.2 The infrared spectrum of Titan as recorded by *Voyager 1*. Compounds responsible for the more prominent features in the spectrum are identified. The horizontal scale is wavenumber, the reciprocal of wavelength, so wavelength increase from right to left. This spectrum covers the range between 7 and 60 microns. Adapted from an illustration in 'Titan' by A. Coustenis and R. Lorenz in *Encyclopedia of the Solar System*, eds P. Weisman *et al.*, Academic Press, 1999.

Figure 3.3 Variations with latitude on Titan of the concentrations of different gases as measured by *Voyager 1* and analysed by Athena Coustenis and Bruno Bézard. Several of the compounds are noticeably more abundant at high northern latitudes where the polar hood was observed. Adapted from a diagram by Coustenis and Bézard published in *Icarus*, vol. **115**, p. 126 (1995).



Something in the air

Methane, made of one carbon atom and four hydrogen atoms, splits when it receives a photon of ultraviolet light. It breaks up into one or two loose hydrogen atoms and a fragment containing a carbon atom and two or three hydrogens. These loose atoms and fragments (called 'radicals') are fiercely reactive. They quickly combine with whatever they find but their couplings are brief, each being followed by a rapid divorce until their promiscuity is halted because all of their energy if dissipated. Most of the hydrogen atoms ultimately combine with each other in pairs, forming molecular hydrogen. Light-weight hydrogen molecules move rapidly enough to escape from Titan's weak gravity and fly free into space over a period of 1-10 million years. They do not, however, have enough energy to escape so quickly from mighty Saturn's gravity, so they collect in a doughnut shaped cloud around Saturn, forming a 'torus' enclosing Titan's orbit. The fragments of hydrocarbons combine in countless different ways to form more complicated organic compounds. For example, two methane molecules each missing a hydrogen may combine to form ethane.

Extract 4

All this photochemistry is interesting for chemists but it only goes half way to addressing theories on the origins of life, a connection which is often used to justify the exploration of Titan. The problems is this. Barring some tiny traces we'll discuss later, Titan's atmosphere is bereft of oxygen or oxygen-containing compounds like water. This is a crucial point for two reasons. First, the chemistry that sustains life is mediated by liquid water; it requires liquid water as a solvent. Second, virtually every organic molecule of biochemical interest contains some oxygen – that means every amino acid, and therefore every protein and enzyme, every sugar, every fatty acid, and DNA itself. If there is no way to incorporate

oxygen into the chemical chain in the atmosphere (and there isn't) then the organics that make Titan's atmosphere such a murky soup will only ever be sterile hydrocarbons and nitriles. However, all it might take to transform this stuff that falls manna-like from the sky is exposure to molten ice. Just add water, as if it were a packet soup.

The outer layers of the young Titan would have been essentially an ammonia solution but in the present era Titan's surface is far too cold for liquid water. At 180 degrees below freezing point, all water is rock solid. The technicalities that might allow liquid water to persist on Mars – thin films of water bound onto mineral grains and the antifreeze effects of salts – aren't enough to save Titan's water from the cold. Even the potent antifreeze ammonia doesn't open a loophole wide enough for liquid water. Thus, the big picture says that Titan has no liquid water and the chemistry stops when material from the atmosphere rains out onto the surface.

The same logic applied to Earth says that Earth is too cold for molten silicates – in other words, liquid rock. Happily, this is true by and large. But in particular regions at particular times conditions far out of the ordinary can take hold. Episodes of volcanism illustrate the point. Material erupted from Earth's interior as hot liquid can trigger chemical processes that do not usually occur. Most mineral deposits of economic value are formed by hydrothermal systems that sweat particular metals out of the rock and deposit them in concentrated form elsewhere.

An ammonia-rich ocean may still exist, deep in Titan's interior. A. Dominic Fortes at Imperial College London noted that conditions beneath Titan's surface may even today allow life to persist if it ever evolved when the ocean was exposed on the surface and able to interact with methane, organics, sunlight and incoming meteorites. Life down there would not be thrill-a-minute but the temperature, pressure and chemical environment would not be hostile to some of the more hardy and persistent life forms on Earth.

It is certain that comets and meteorites have slammed into Titan from time to time. When a very large meteorite or asteroid strikes a planetary body, most of its energy goes into a huge explosion, which excavates a crater and may throw material high into the atmosphere, or even into space. A proportion of the energy carried by the impactor goes into heating the target rock, some of which may melt. The larger the crater the greater the volume of rock that melts in the impact. One of the biggest impact structures on Earth (but sadly without an obvious crater shape) is the two-billion-year old Sudbury structure in Canada. The huge pool of molten rock lasted long enough for metal blobs in the rock to settle out at the bottom of the melt pool. Today these blobs, now long solidified, are one of the largest nickel deposits in the world.

Reid Thompson at Cornell University and his advisor Carl Sagan realised in 1991 that this kind of environment could exist on Titan. The ice would be melted in large impacts and would take thousands of years or longer to freeze again. So, locked up in Titan's icy surface, there is likely to be a plethora of molecules formed by the interactions of water with the ubiquitous hydrocarbons and nitriles. Amino acids are sure to be there but who knows how complicated the chemistry can get? A sophisticated surface-sampling mission to Titan in the future will be needed to find out.

Extract 5

The Cassini-Huygens mission

Nearly all our hopes for uncovering Titan's secrets are pinned on a robotic monster, the size and weight of a large dinosaur. With 7.5 miles of wire for a nervous system, it weights 5.5 tonnes and is 6.8 m long. The multibillion dollar enterprise to get this surrogate explorer to Saturn spans nearly two decades and involves 15 countries. By the end of its mission, around 7100 work-years will have been spent on it.

Genesis of Cassini

A mission involving an orbiter around Saturn and one or more atmosphere probes was already being thought about in the later 1970s, just as the *Voyagers* started on their way. It was the logical next step and the planning of the corresponding mission for Jupiter, what was to become *Galileo*, had already started. One of the earliest NASA ideas was 'SO2P' – a Saturn orbiter with two probes, one for Saturn (which would make a useful comparison with the results of the *Galileo* probe into Jupiter's atmosphere) and one for Titan. But realistic plans for a Titan probe couldn't be made until after the results of the *Voyager* flyby established how thick Titan's atmosphere was, information crucial for the design of heat shields and parachutes.

Thinking started in earnest soon after the *Voyager* encounters. The original idea, as proposed by a consortium of European scientists to the European Space Agency ESA, was that NASA would supply a duplicate of its *Galileo* probe and Europe would build an orbiter – the mother ship that would deliver the probe. A joint NASA-ESA study team was set up. The NASA study scientist was JPL's Wes Huntress, who later went on to be the NASA Associate Administrator for Space Science – the top job in Space Science in the USA; in ESA Jean-Pierre Lebreton, an affable French scientist, led the study.

Soon it was decided to reverse the agency roles: ESA would provide the probe, while NASA would provide the orbiter. The orbiter would be based around a new 'bus', called *Mariner Mark II*, after the successful *Mariner* series of the late 1960s and early 1970s. Spacecraft engineers use the term 'bus' as auto engineers might use the word chassis – a basic structure around which different variants of the same basic vehicle could be easily built, without 'reinventing the wheel'. For a spacecraft, this does not mean only the mechanical structure but also many of the complex systems, like propulsion, power, attitude control (pointing) and so on.

The first two spacecraft in this series, which (as they all do) promised to reduce costs, were to be *Cassini* and *CRAF (Comet Rendezvous and Asteroid Flyby)* which, as its name suggests, was to explore some of the smaller bodies of the solar system. One of the *CRAF*'s distinctive highlights was that it would fire a penetrator – like a large instrumented harpoon – into the comet to measure some of its properties *in situ*, while the mother ship orbited overhead. Both *CRAF* and *Cassini* would have scan platforms with cameras and spectrometers, and instruments to study dust, gas and plasma in their respective environments.

New challenges for Europe

In 1984, an assessment study was performed ESA – a preliminary investigation to identify the principal technical challenges. A planetary probe was a new venture for ESA, whose only interplanetary spacecraft was *Giotto*, at that time on its way to fly past Comet Halley in 1986. The fiery heat of entry into a planet's atmosphere was a new challenge and such things as parachutes were also a novelty for engineers who were more used to solar panels and thruster nozzles.

The entry protection was a delicate issue. It would be crucial to the success of the project but was also tied up with strategic concerns. The sciences of hypersonic aerodynamics and aerothermodynamics are intimately linked – as is rocketry – with the intercontinental delivery of nuclear weapons. Within the ESA member states, and remember that ESA is an agency with an exclusively peaceful mandate, only the industries of Britain and France had experience with such materials, and much of that experience would have to be kept secret.

In the assessment study phase, engineers in ESA and industry developed impressive and imaginative ideas. One study examined spacecraft autonomy. A Titan probe would be too far away for ground control to respond to problems. The study mapped out ways – essentially implemented only later in NASA's 1998 *Deep Space 1* mission – whereby onboard computers could respond to an unknown environment or equipment failures. For example, if a battery failed, reducing the total amount of energy available for the mission, it would shut down some lower-priority experiments to save energy for the most important near-surface studies. Other creative concepts addressed the aerothermodynamics, advocating the use of exotic materials like carbon composites and beryllium, a marvellously light, hard metal with a high melting point but one that it notoriously difficult to work with.

After the assessment study, ESA then went on to fund a 'Phase A' study – still only paper, but with the aim of developing a credible overall design, with a reliable cost estimate. In the absence of commitment from the US, the start of the Phase A study was prudently delayed a year to November 1987, to synchronise NASA's schedule with ESA's. ESA's space science budget is fairly small and so there are few missions under development at one time. The Titan probe was in competition with other candidates for 'M1', ESA's first 'medium-class' mission, including *Vesta* (a mission to rendezvous with the asteroid of that name) and astronomy satellites. The Titan probe was selected in November 1988 pending NASA approval of *Cassini* and was named *Huygens*.

Meanwhile, in the USA, *Cassini* had bubbled up to the top of NASA's to-do list and the US Congress approved the start of the project in 1989. NASA and ESA jointly solicited proposals for scientific investigations. Americans and Europeans would compete to provide instruments for both the Titan probe and the orbiter. ESA also invited proposals from European industry to build the probe itself, which would cost around 300 million accounting units (an ESA 'currency', approximately equal to one US dollar or euro). Things were now getting serious. In October 1990, the selected experimental payloads for *Huygens* were announced.

In many respects the most important experiment, and certainly the most massive, was the gas chromatograph/mass spectrometer (GCMS). This 18-kg experiment will analyse both the composition of atmospheric gases and the material delivered from the aerosol collector and pyrolyser. Because Titan's atmosphere is so chemically complex, a simple mass spectrometer (as flown on *Galileo*, for example) would not be able to distinguish the many chemicals that might have the same molecular mass. The GCMS identifies compounds not only by molecular mass but also by their affinity to special coatings inside thin tubes, along which different compounds take different times to travel. The combination of the two techniques is very powerful. The GCMS has dozens of valves, high voltage components, delicate filaments and other design challenges. It was to be tackled by a team led by Hasso Niemann of NASA's Goddard Spaceflight Center. Hasso had built the mass spectrometer on *Galileo* as well as a number of similar instruments.

The second US-led experiment was the descent imager/spectral radiometer (DISR). As well as taking pictures looking down and sideways, this instrument will measure the spectra of sunlight filtering through the haze and reflected up from the ground. It will also measure the sunlight scattered around the Sun (the aureole), which will yield data about properties of the haze. The team is led by Marty Tomasko of the University of Arizona. The CCD detector was provided by Uwe Keller of the Max-Planck Institute for Aeronomie in Katlenburg-Lindau in Germany, who built the camera on ESA's *Giotto* spacecraft to Halley's Comet. The infrared spectrometer came from Michel Combes of the Paris Observatory working with other spectroscopic and imaging experts from the USA, France and Germany.

One of the most conceptually simple experiments on the probe is the Doppler wind experiment (DWE). The only hardware associated with it on the probe is an ultrastable oscillator on one of the two channels of the probe radio link. By measuring the change in frequency of the signal received, aided by a second reference oscillator on the orbiter, the Doppler shift due to the probe's motion can be derived. Most of this Doppler shift is the result of the orbiter's rapid approach towards Titan and some is due to the probe's vertical descent. These two components can be calculated and removed. Any remaining component of the probe's motion is due to wind. The DWE team is led by Mike Bird of the Radio Astronomy Institute at the University of Bonn in Germany. German, Italian and American colleagues rounded out his team.

The aerosol collector/pyrolyser (ACP) is a novel instrument that sucks air through a tiny filter held out in front of the probe, trapping aerosol particles. The filter is then pulled inside the instrument and heated in a set of temperature-controlled ovens. The aerosol material breaks down under heating and the gaseous products released are transferred through a pipe to the GCMS instrument for analysis. The development effort is led by Guy Israel of the Service d'Aeronomie in Paris. Guy had built a similar, but simpler, instrument for the

Soviet *VEGA* mission to Venus some years before. His team included Austrian electronics providers and collaborators from the GCMS team in the USA. Although conceptually simple, the instrument posed severe technical challenges and the two filters originally proposed were reduced to only one.

The *Huygens* atmospheric structure instrument (HASI – originally just ASI until the Italian Space Agency was formed with those same initials, requiring 'Huygens' to be added) will measure the basic properties of the atmosphere during the entry and descent. An organisational nightmare, tackled with aplomb by Italian Principal Investigator (PI) Marcello Fulchignoni, this experiment brought together temperature sensors and central electronics from Italy with pressure sensors from Finland, accelerometers assembled in the UK and electrical sensors and electronics built by a consortium with members in ESTEC, Austria and Spain. For good measure, scientific expertise from the USA, Germany, Israel, France and Norway added to the team. There is probably scope for a joke somewhere: 'Heaven is where the Bankers are Swiss, the Chefs from France, the pressure sensors from Finland…'.

Lastly (and it always gets mentioned last) is the surface science package (SSP). This is a collections of small sensors (similar to HASI in that respect) led by John Zarnecki, formerly of the University of Kent at Canterbury, now at the Open University in the UK. As well as an accelerometer and penetrometer to characterise the impact with the surface, the package includes tilt sensors to measure the probe's orientation and bobbing in any waves; the SSP team was unique in including an oceanographer. An ingenious refractometer will measure the ethane/methane composition of an ocean, with additional information provided by thermal, density and electrical permittivity sensors. An acoustic sensor would act as a sonar to measure the depth in a liquid. While most of the equipment was built in the UK, the acoustic sensor was provided by ESTEC and thermal sensors were developed in part in Poland. An X-ray fluorescence sensor, which was to have been provided by the USA, had to be dropped for budgetary reasons.

Although no radar science instrument was selected, a capability for HASI to digest the signal from the probe's engineering altimeter was added later. The option to include a nephelometer, which would measure the light scattering by haze particles using a laser and a small mirror (actually it was the flight spare from the *Galileo* probe), was also studied but determined not to be worth the effort (and the 5 kg) since DISR would measure essentially the same things.

The contributors of experiments on *Huygens* are distributed fairly uniformly across Europe and the USA. It is interesting to note that being the PI for an instrument is more an organisational job than a scientific one. Of the *Huygens* PIs only Tomasko has worked on Titan before.

Extract 6

Mission Plan

The plan in the Phase A study was to launch *Cassini* in April 1996. After flybys of Earth, an asteroid and Jupiter, *Cassini* would arrive at Saturn in 2002. The details of this plan (and crucially, the dates) would change several times in the years to come but, because *Cassini* is so massive that even NASA's largest launcher, the Titan IV, could not project *Cassini* directly to Saturn, most of the incarnations of the mission incorporated an Earth encounter to slingshot *Cassini* out of the inner solar system and a gentle nudge by Jupiter a couple of years before Saturn arrival. In the end, Saturn arrival would occur on the 1st July 2004.

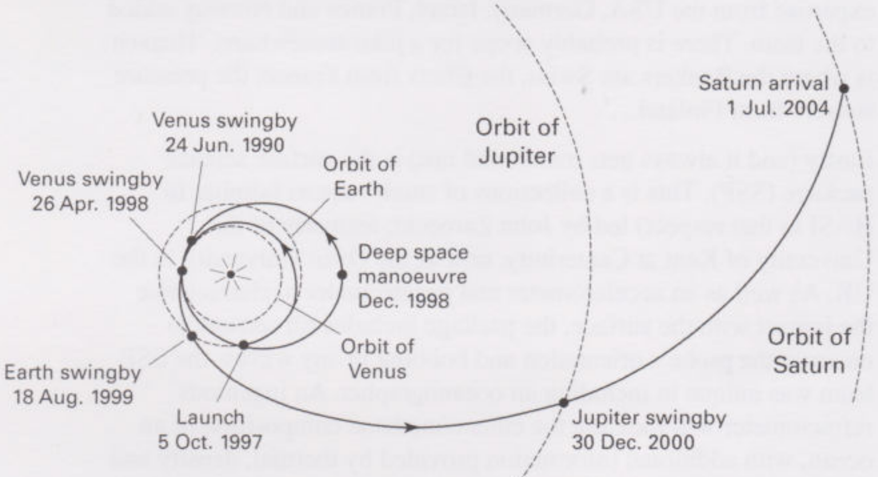


Figure 6.4 The interplanetary flight path of the *Cassini* spacecraft, beginning with launch on the 15th of October. There were gravity assist flybys at Venus on the 26th of April 1998 and the 24th of June 1999, at Earth on the 18th of August 1999 and at Jupiter on the 30th of December 2000.

Cassini's engine will brake it into a long comet-like orbit around Saturn and the craft will climb, ever more slowly, out to several million km. At this point, a couple of months after arrival, its engines fire once again to raise its orbit's point of closest approach to Saturn, the 'periapsis', so that *Cassini* will never again to get closer than about three Saturn radii. This periapsis raise manoeuvre (PRM) also sets *Cassini* up to fly towards Titan.

The first order of business for probe delivery is to get the batteries working. The lithium cells will have built up a thin layer of oxide on their electrodes, which impedes the flow of current (and for that matter keeps the batteries fresh for all those years). By drawing a large current for a short time this layer is burnt off, enabling strong and consistent battery performance afterwards. After the final upload of commands and sequences from the ground, explosive bolts will fire and release the probe which is pushed away by springs. A clever mechanism, called the spin eject device, gives a gentle stabilising spin to the probe as it is pushed away. As the springs push, special low-force electrical connectors come unstuck, literally cutting *Huygens'* umbilical cord to its mother ship.

From this point on, *Huygens* is on its own, like a clockwork toy. It will also get colder than it has ever been or ever will be during its mission. All its systems are shut down, except for three quartz

clocks, set to wake it up 18 minutes before it hits Titan's atmosphere. The probe would in fact get too cold for crucial components like the batteries to survive, so it is kept warm with a couple of dozen pellets of plutonium supplying just enough heat to enable *Huygens* to be safely resuscitated. It will coast, asleep and gently spinning, for 22 days, during which time Titan will sweep around Saturn one and a half times to appear underneath *Huygens* at their appointed rendezvous, initially planned for the 27th November 2004 but now set for the 14th January 2005, three Titan orbits later.

Cassini will in the meantime reorient itself and make another burn of its engines. This burn slows *Cassini* down with respect to *Huygens*, so that *Cassini* arrives several hours later (the so-called Orbiter Delay Time, ODT) and can act as a radio relay. It also displaces *Cassini* to one side, so that the orbiter flies by, rather than into, Titan.

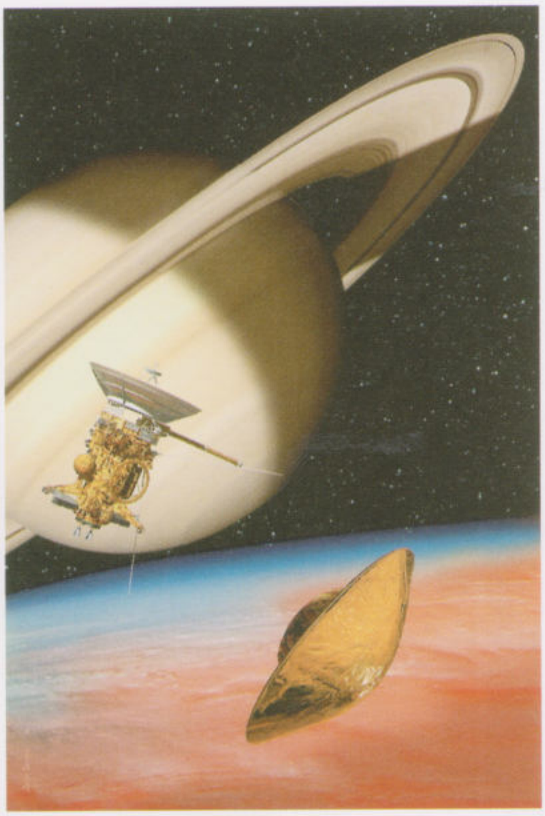


Figure 6.5 An impression by ESA artist David Ducros of the *Huygens* probe shortly after release from the *Cassini* spacecraft. When the event actually takes place, Titan will be farther away than depicted here. The artist has also been unable to resist the temptation to show the rings ajar; in reality, they would appear edge-on. Artwork: ESA.

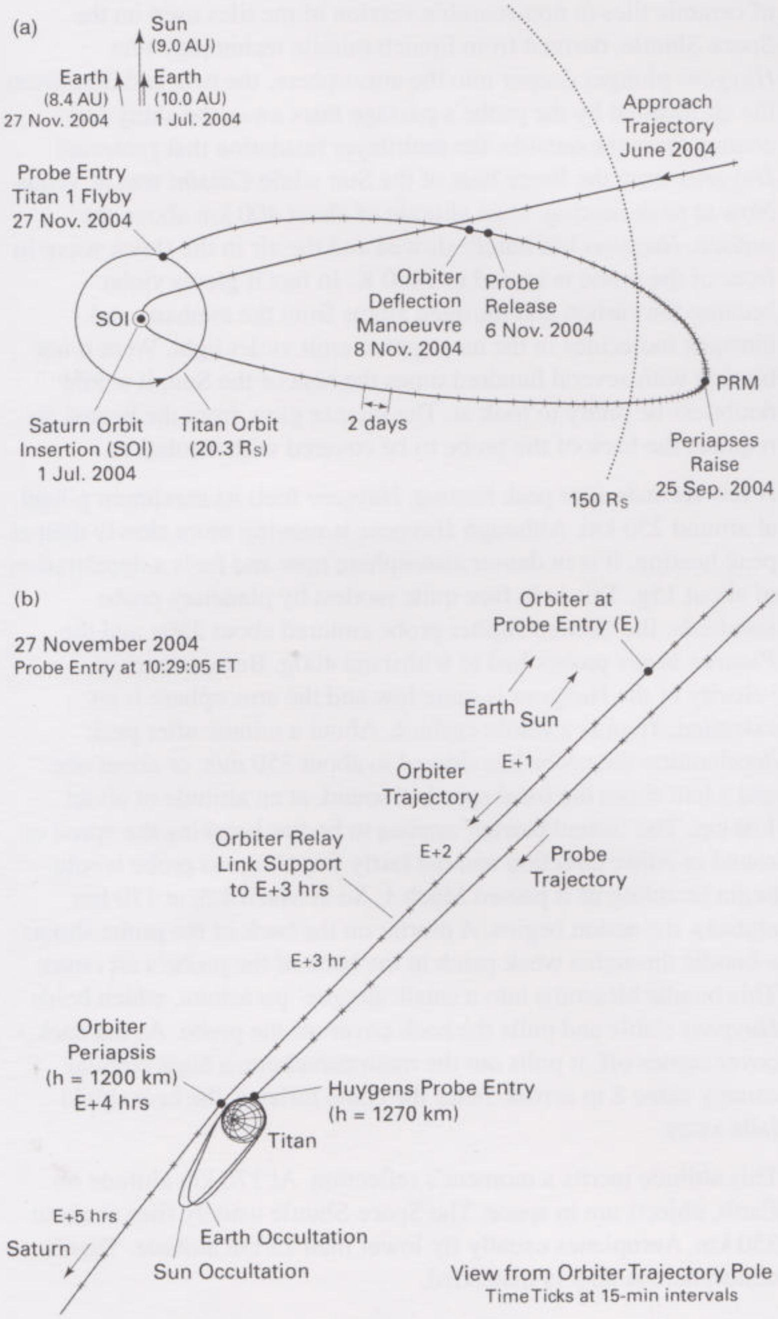


Figure 6.6 (a) The original plan for *Cassini*'s trajectory between its initial arrival at the Saturn system in June 2004 and the arrival of the *Huygens* probe at Titan in November 2004. In the new plan, the first orbit is smaller, encountering Titan in October 2004. *Cassini* makes 2_ orbits before the probe is released. (b) Details of the original trajectories of *Cassini* and *Huygens* just before and during the probe's entry into Titan's atmosphere and descent to the surface. In the revised plan the orbiter is much farther away (periapsis >5000 km).

Cover me – I'm going in

Huygens wakes up a few thousand kilometres above Titan, travelling at a speed of about 6 km/s. The computers boot up and various instruments are turned on. Most electronic systems perform best at constant temperature. A particular example is the ultrastable oscillator on the Doppler wind experiment, which takes a short while to settle down to its most stable frequency.

The first science *Huygens* does begins at about 900 km altitude, when the thin atmosphere begins to tug noticeably at the craft. Noticeably means the detection threshold of the accelerometers of the HASI experiment: a few millionths of Earth's gravity (i.e. 10^{-6}g), when the atmosphere is about a billion times thinner than at Titan's surface. *Huygens* is slowed by the drag on a blunt conical heat shield, 2.7m across. The shield is protected by a 1-2-cm-thick layer of ceramic tiles (a non-reusable version of the tiles used on the Space Shuttle, derived from French missile technology.) As *Huygens* plunges deeper into the atmosphere, the heat and drag from the air tortured by the probe's passage tears away the shiny foil coating from the outside, the multilayer insulation that protected *Huygens* from the fierce heat of the Sun while *Cassini* was at Venus. Now at peak heating, at an altitude of about 400 km above the surface, *Huygens* has barely slowed and the air in the shock wave in front of the probe is heated to 1400 K. In fact it glows violet because the carbon and nitrogen atoms from the methane and nitrogen molecules in the air together emit violet light. Were it not blazing with several hundred times the heat of the Sun, it would doubtless be pretty to look at. The intense glow from the heated air requires the back of the probe to be covered with insulation.

A few seconds after peak heating, *Huygens* feels its maximum g-load at around 250 km. Although *Huygens* is moving more slowly than at peak heating, it is in denser atmosphere now and feels a deceleration of about 15g. This is in face quite modest by planetary probe standards: the *Galileo* Jupiter probe endured about 250g and the *Pioneer Venus* probes had to withstand 400g. Because arrival velocity of the H*uygens* is quite low and the atmosphere is so extended, Titan is a gentle cushion. About a minute after peak deceleration the probe has slowed to about 350 m/s, or about one and a half times the local speed of sound, at an altitude of about 170 km. The 'sound barrier' applies to bodies breaking the speed of sound in either direction and the fairly flat *Huygens* probe would begin tumbling as it passed Mach 1. So at Mach 1.5, at 170 km altitude, the action begins. A mortar on the back of the probe shoots a bundle through a weak patch in the back of the probe's aft cover. This bundle blossoms into a small 'drogue' parachute, which holds *Huygens* stable and pulls the back cover off the probe. As the back cover comes off, it pulls out the main parachute, a huge circular canopy some 8 m across. After the chute inflates, the heat shield falls away.

This altitude merits a moment's reflection. At 170 km altitude on Earth, objects are in space. The Space Shuttle usually flies at about 350 km. Aeroplanes usually fly lower than 15 km altitude. Titan's atmosphere is hugely distended.

As the main chute opens, the probe decelerates quickly to a steady descent rate of about 50 m/s. A cover that has protected DISR from material coming off the heat shield is released and flies off on springs. Similarly, caps on the inlet pipes of the GCMS are broken off by explosive actuators and allow gas sampling to begin. Two small arms bearing HASI's electrical field sensors swing out and lock into position. Small vanes mounted around the edge of the probe keep up a slow spin under the parachute to pan DISR's camera around in all directions; the line to the parachute has a swivel to prevent it twisting up.

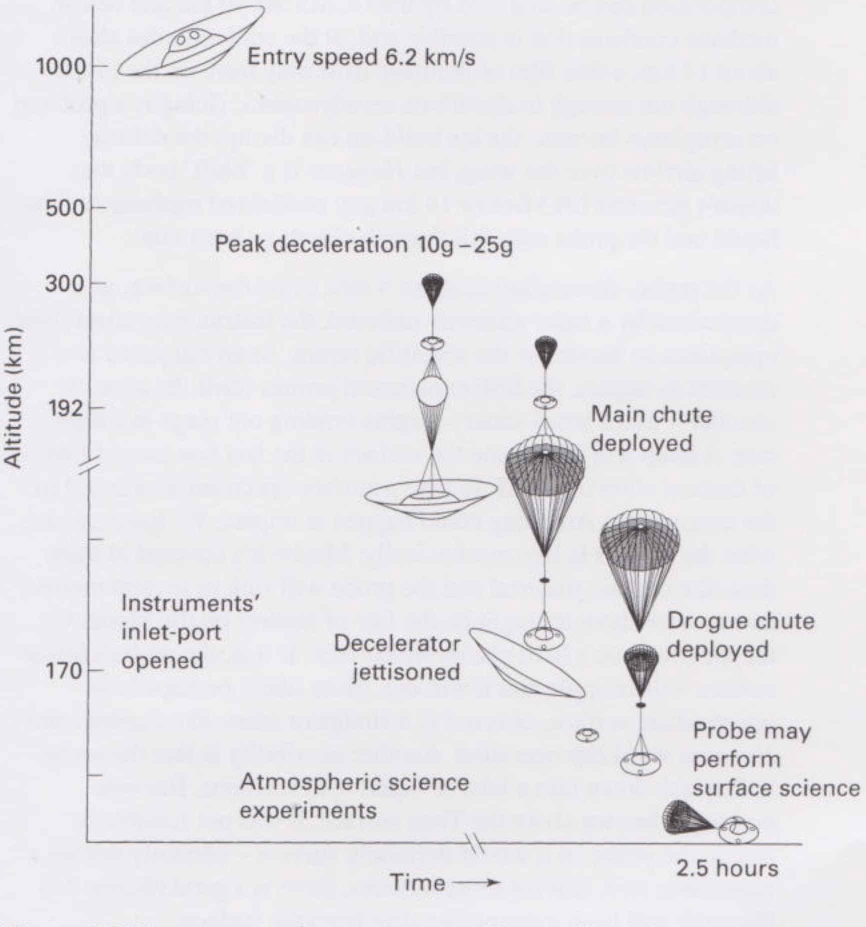


Figure 6.7 The sequence of events between *Huygens*' initial encounter with Titan's atmosphere and its impact on the moon's surface. As the probe slows down, a small parachute will be released, which will then deploy the main parachute. Once the main parachute is fully open, the decelerator shield will be jettisoned and the probe will drift downwards. About 40 km above the surface the main parachute will be jettisoned and a smaller drogue chute will carry the probe the remaining distance. Scientific data will be transmitted continuously to the *Cassini* orbiter during the 2.5-h descent and relayed to Earth later. If the probe survives its impact at about 10miles per hour, it may continue to transmit data to *Cassini* for up to a further 30 min.

The size of the main parachute was dictated by the need to pull the probe away from the heat shield safely. But if the probe continued to descend under the main parachute, it would take some 5-8 h to reach the ground, by which time, the *Cassini* orbiter would have disappeared out of sight before the surface data could be taken. So after 10 minutes the final explosive bolts are fired to release the main parachute. A smaller stabiliser parachute allows the probe to descend rather more quickly.

The ACP experiment sucks atmosphere in through a filter, trapping aerosol particles. A small oven breaks them down so their composition can be analysed by the GCMS. At 40 km and below, methane condensation is possible and, at the cold altitudes above about 14 km, a thin film of methane frost may form on the probe, although not enough to disturb its aerodynamics. (Icing is a problem on aeroplanes because the ice build-up can disrupt the delicate lifting airflow over the wing, but *Huygens* is a 'bluff' body that doesn't generate lift.) Below 14 km any condensed methane will be liquid and the probe may fall through clouds or even rain.

As the probe, descending at about 5 m/s, nears the surface, as determined by a radar altimeter onboard, the instruments adapt their operations to maximise the scientific return. At an estimated two minutes to impact, the SSP experiment primes itself. Its acoustic sounder – like a small sonar – begins sending out pings at a high rate. A lamp will illuminate the surface in the last few tens of metres of descent allowing DISR to take a surface spectrum unaffected by the atmosphere. Anything could happen at impact. We have no idea what the surface is like mechanically. Maybe it's covered in thick dust-like organic material and the probe will sink in several metres, as was once thought might be the fate of landers on the Moon. Or maybe it will be a hard glazed ice surface. If it is, the probe's lower surface will crumple and it will die. More likely perhaps is an intermediate surface, covered in a sludge or snow-like deposit, and *Huygens* will keep operating. Another possibility is that the probe will splash down into a lake of liquid hydrocarbons. Because nothing is known about the Titan surface, it was not feasible to design the probe so it would definitely survive – certainly not for a reasonable cost. But for most surfaces, there is a good chance that *Huygens* will keep transmitting data from the surface.

Once concern was whether the probe would float and, if so, whether it would it do so with the parts of the probe requiring access to the liquid (like the Surface Science Package) below the 'waterline' but with the camera above it? It will. In fact the period with which it bobs up and down can be used to infer the density and the composition of the liquid.

If nothing breaks at impact, the probe will continue to take data but it should be borne in mind that, unlike *Mars Pathfinder*, the probe is not designed for operations after it has landed. There will be pictures but they will be limited to wherever the camera points when it comes to rest, whether at a rock, or blank sky, or maybe even the underside of the parachute. There's a chance that some surface material may be vaporised by the heated inlet of GCMS so that it can analysed. In a liquid the acoustic sounder on the probe may pick up an echo from the bottom and hence determine its depth.

Ultimately, one of several things will happen. The batteries will run out or, as the probe cools down, something – perhaps the batteries, perhaps the radio transmitter – will stop working. *Cassini* flying over-head will disappear out of sight, although somewhat before then the accuracy with which the antenna on *Huygens* is pointed at *Cassini* will become so poor that the data link will die out.

After listening out for the probe, the *Cassini* orbiter will slew around to take pictures and spectra of the areas as near the landing site and as close to the *Huygens* descent time as possible. These data will be key in understanding future Titan observations from *Cassini*. Several hours after the Titan flyby and the end of the probe mission, *Cassini* will slew over to point at Earth and will disgorge its hard-won load of data from the *Huygens* mission.

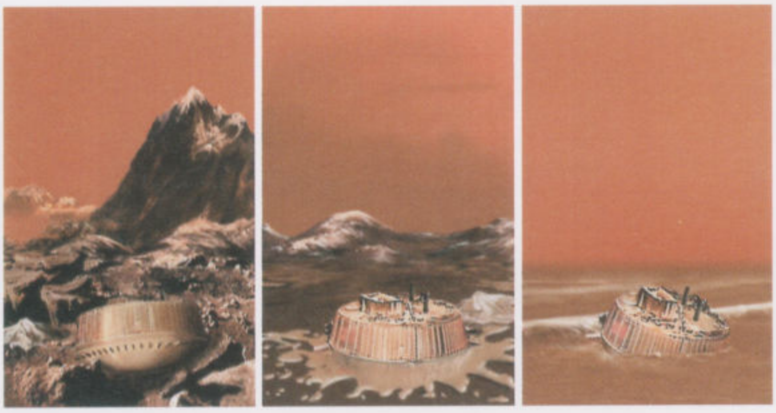


Figure 6.8 An artist's impression of three possible scenarios for the landing of the *Huygens* probe on Titan. From left to right: the 'crunch' (on solid terrain), the 'squelch' (in slush) and the 'splash' (in liquid). Artwork by James Garry.

Astronomers have detected over 50 giant planets outside our solar system and made remarkable progress in determining their properties. But the real prize would be an Earth-like planet that could harbour life

Extrasolar planets

Andrew Collier Cameron

AS RECENTLY as five years ago, the search for planetary systems beyond our own was the subject of a few painstaking surveys by a small but dedicated band of planet-hunters. Their goal was to find extrasolar counterparts of our own giant Jupiter, which circles the Sun once every 11 years. To do this they relied on the fact that the star around which any planet orbits also moves: for instance, the Sun circles its common centre of gravity with Jupiter at the leisurely pace of 12 metres per second. The aim of these early programmes therefore, was to develop techniques that could detect the periodic changes in the Doppler shift of the light from a star as it wobbled back and forth in response to the gravitational tug of an unseen Jupiter-like companion (figure 1).

Several teams achieved the required precision by the early 1990s. These efforts included a regular monitoring programme of 120 nearby stars that was started at the Lick Observatory in California in the late 1980s by Geoff Marcy and Paul Butler of San Francisco State University. However, the computational overheads needed to analyse the results were so great that much of their data remained unstudied in an archive. This strategy made good sense: by the time any candidate Jupiters had completed enough of their orbits to be clearly identifiable, processor speeds would have increased to the point where the data analysis could be carried out far more quickly. Other groups at the University of Victoria in Canada and the University of Texas at Austin adopted similar philosophies.

When Michel Mayor and Didier Queloz of the Geneva Observatory announced the first discovery of stellar reflex motion due to a planetary body in 1995, its form was so unexpected that it demanded independent confirmation. Their data showed the solar-like star 51 Pegasi to be wobbling back and forth at 56 $\mathrm{m\,s^{-1}}$, completing one orbit every 4.2 days. The only plausible explanation for this wobble was the presence of an unseen body with at least half the mass of Jupiter in an orbit with a radius of 0.05 astronomical units (AU). An astronomical unit is defined as the mean distance between the Earth and the Sun, which is around 1.5×10^8 km.

Fortunately Marcy and Butler had been monitoring the same star for several years, and confirmed almost immediately that the wobble was present in their data as well. At the same time, they found similar signatures of close-orbiting giant planets for three other stars: tau Bootis, 55 Cancri and upsilon Andromedae. A spate of similar discoveries followed. New monitoring programmes were established, and the total number of stars that are currently under observation stands at about 1000. Among these, the tally of nearby stars known to possess at least one planet has risen to over 50. Several of these stars show extra, slower wobbles superimposed on the main signal, which betray the presence of one or more additional planets in larger orbits. Most notable among these is upsilon Andromedae with a family of three giant planets.


PPARC

Exoplanetary systems have inspired new theories of planet formation

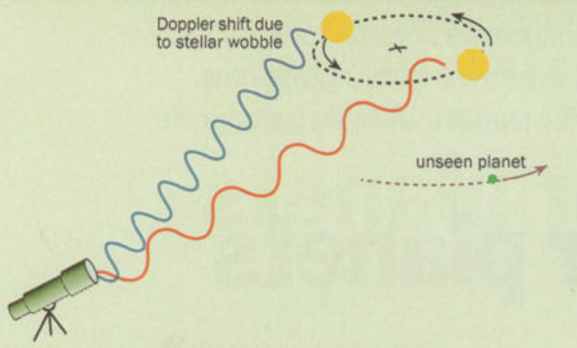
Last year saw the record for the lowest mass extrasolar planet being broken several times. Mayor currently holds the record for detecting a planet weighing just 0.16 times the mass of Jupiter that is orbiting around HD 83443, a bright star in the constellation Vela some 141 light-years from Earth.

Meanwhile in August 2000, a team led by Bill Cochran of the University of Texas at Austin announced that it had detected the closest extrasolar planet to Earth. At a distance of just 10.5 light-years away, the star epsilon Eridani shows evidence of a seven-year wobble, although this result still needs to be confirmed.

Reflex orbit

The period and size of a star's wobble encode important information about the planet's mass. Usually it is only possible to determine a lower limit on the mass, because for most systems astronomers cannot measure the tilt of the orbit relative to the line of sight. Assuming that the orbit is edge-on to the line of sight gives the smallest possible value of the planet's mass. This is not as serious a problem as it might

1 Stellar wobble



A star and planet orbit around their common centre of mass. Although the planet is so faint that it cannot be seen, the star's reflex motion Doppler-shifts the starlight alternately to longer and shorter wavelengths that can be detected using high-precision spectroscopy.

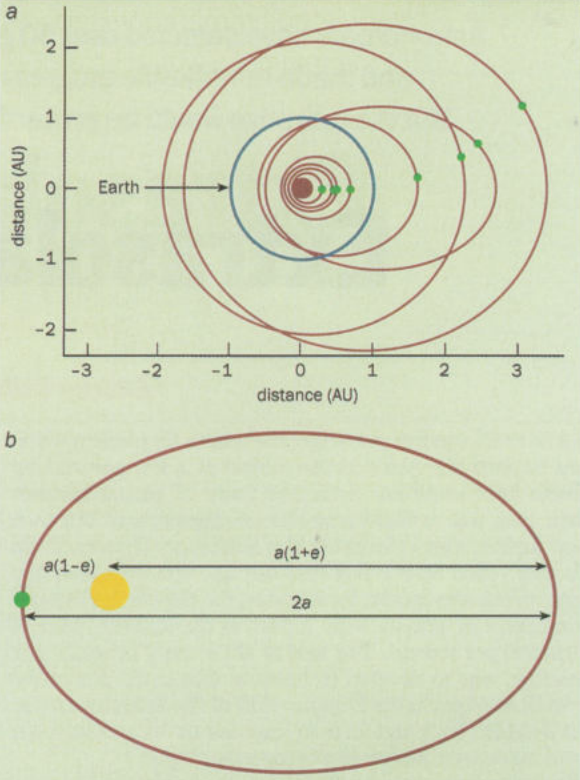
appear. If the orbital axes of the planetary systems are oriented randomly in space, there is a natural statistical tendency for us to see many more orbits edge-on rather than face-on. (Try throwing a handful of coins in the air to convince yourself of this.) All the planets discovered so far are giants with masses ranging from slightly less than the mass of Saturn, which is 95 times heavier than Earth, to a dozen times the mass of Jupiter. Jupiter is the most massive planet in our solar system and is 320 times heavier than the Earth.

A planet in a circular orbit around its star produces a symmetric wobble that varies sinusoidally with time. The dozen or so known planets with periods of less than a week – the so-called hot Jupiters, deemed hot because they are substantially closer to their stars than our Jupiter is to the Sun – have orbits that are nearly circular, as expected (figure 2). The reason is that a close-orbiting planet in a highly elliptical orbit produces tides on the star, which move the star's centre of gravity in such a way that the orbit gradually evolves into its lowest energy state, a circular orbit. The greater the distance between the planet and the star, the smaller the tidal effects and the longer it takes for the orbit to become circular. For planets that have orbital distances larger than a small fraction of an astronomical unit, the orbit will remain elliptical for longer than the star's lifetime.

A planet in an elliptical orbit speeds up when it is close to the star and slows down when it is further away from it, giving the star a characteristic lopsided wobble. Astronomers have used this idea to deduce that the orbits of the longer-period planets are much more eccentric (i.e. more elliptical), in stark contrast to the near-circular orbits of the planets in our solar system. Indeed, nobody has yet found a well behaved Jupiter-like planet in a circular orbit with a radius of several astronomical units.

The high masses and short orbital periods of the planets discovered so far makes them easier to detect than conventional Jupiter-like planets with long periods. This is because the speed, v, of the parent star in its reflex orbit is given by the relation $v = 12(M_P/M_J)(5.2M_{Sun}/aM_{star})^{1/2}$ m s^{-1} where M_P, M_J, M_{Sun}, M_{star} are the masses of the extrasolar planet, Jupiter, the Sun and the star, respectively, a is the orbital distance of the planet and 5.2 AU is Jupiter's distance from the Sun. This favours the discovery of massive planets in close short-period orbits, particularly as most of the monitoring programmes have only been running for a few years. However, it does not explain why so many of the orbits are so much more eccentric than those found in our solar system. To understand these differences, we need to look at the conditions under which planetary systems form.

2 Exoplanet orbits



(*a*) A selection of the orbits inferred for several extrasolar planets (drawn to scale) together with the Earth's orbit (blue circle). The "hot Jupiters" have near-circular orbits and are buried at the centre of the diagram. Exoplanets with orbits more than few tenths of an astronomical unit are in highly elliptical orbits. (*b*) The departure from a circular orbit is quantified by its eccentricity, e. In an eccentric orbit, the star is located at one focus of the elliptical path traced out by the planet. The star–planet distance is $a(1-e)$ at closest approach and $a(1+e)$ at maximum separation, where a is the semi-major axis and e is the eccentricity. Circular orbits have zero eccentricity.

How do Jupiters form?

Theories of giant-planet formation fall into two main categories. The "top-down" approach has planets forming from large-scale perturbations in the flattened, gaseous disc that surrounds a new-born star for the first few million years of its life. Meanwhile, the "bottom-up" approach requires dust grains with ice mantles to clump together to form bodies a few times the mass of the Earth. Once this critical mass is attained, the planet's gravitational pull becomes strong enough for it to accrete large amounts of gas from the disc and grow rapidly into a gas giant.

As the planet grows, however, it causes slow-moving material outside its own orbit to speed up and faster-moving material with smaller orbits to slow down. This tidal effect sweeps

the planet's own orbit relatively free of material (figure 3). Computer simulations show that accretion can only occur along a pair of spiral shocks extending inward and outward from the planet. However, several things can go wrong. For example, the growth process itself may be self-limiting due to a lack of material in the region swept clear by the planet. And in many models, the angular momentum that is inevitably exchanged between the planet and the surrounding disc material causes the planet's orbit to decay, spiralling in to be swallowed by its sun before it can attain a high enough mass.

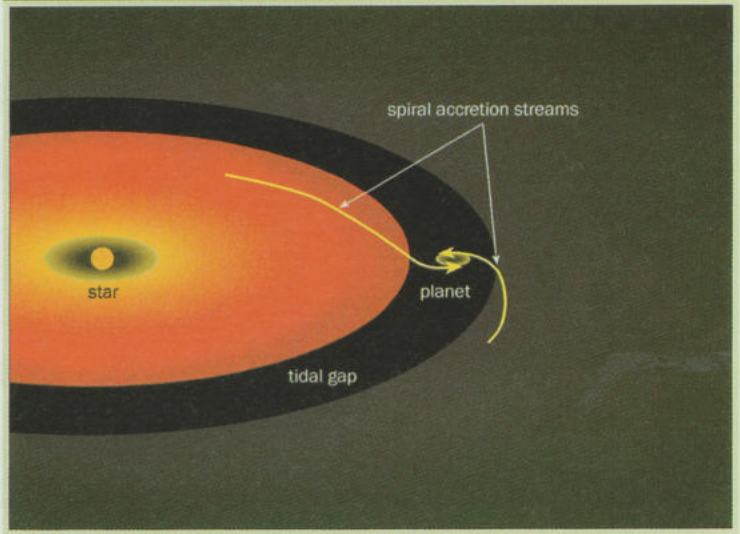
Other models have shown that top-down planet formation occurs in local clumps in the wake of a growing giant planet, producing a system of several giant planets in dynamically unstable orbits. In this case, interactions between the planets may lead to some bodies being ejected from the system altogether, while others are left behind in eccentric orbits.

Tantalizing new evidence emerged recently that top-down formation of planetary-mass bodies may even occur in interstellar space. Maria Zapatero Osorio and co-workers at the Instituto de Astrofisica de Canarias in Spain and the California Institute of Technology in the US recently announced the discovery of several faint objects in the star-forming region around the massive star sigma Orionis (see further reading). The spectra of these "freely floating" objects, which appear unbound to the star, are a good match with theoretical models of gaseous bodies that weigh a few Jupiter masses and are between 1 and 5 million years old. At these very young ages, the objects still shine brightly as they contract by radiating gravitational energy. However, it is not yet clear from the observations whether the free-floaters have formed in isolation or have been ejected from nearby proto-planetary systems.

These new ideas – inspired largely by the properties of the exoplanetary systems discovered so far – paint a much more violent picture of the planet-formation process than is needed to explain our solar system. While the same problems of spiral-in and self-limiting growth were encountered as long ago as the mid-1980s, much of the fine-tuning of the models was carried out under the assumption that the end-product should look like our system, with well behaved giant planets in circular orbits at more or less the distances where they formed.

The orbits of the giant exoplanets suggest that many of the things that can go wrong in building a tidy system like our own, do go wrong elsewhere. Neither top-down nor bottom-up scenarios can produce Jupiter-like planets with four-day orbits, such as the planets around 51 Pegasi, tau Bootis, upsilon Andromedae and the star HD 187123 in the constellation Cygnus. If the cores of these exoplanets formed from rock–ice planetesimals, they must have done so several astronomical units from their stars, accreted their atmospheres, spiralled in and had their migration halted near 0.05 AU by some as yet unknown mechanism. The giants in eccentric orbits between a few tenths and a few AU must have undergone a similar formation and migration history, combined with violent dynamical interactions with other newly formed giants.

3 Planet formation



A high-mass planet forming in a protostellar accretion disc produces a "tidal gap" nearby. Accreting material flows onto the planet along a pair of spiral arms stretching inward and outward from the gap. A net imbalance in the exchange of angular momentum between the planet and the surrounding disc material causes the planet to gradually migrate inwards or, in some models, outwards.

Hints from other worlds

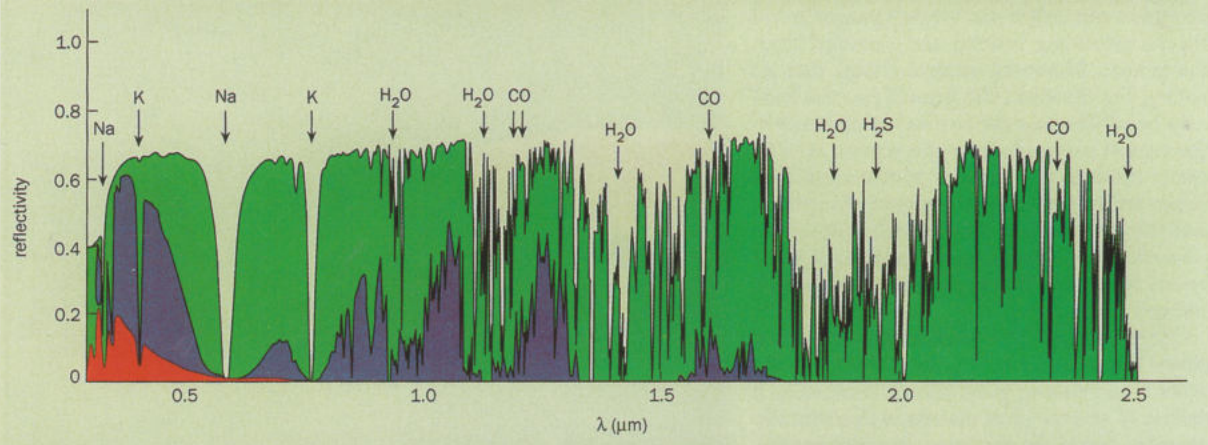
The exoplanets themselves may give some clues about their formation history. The radius of a Jupiter-like planet depends weakly on its total mass and internal composition, and also on its age. Even today, Jupiter is shrinking, radiating 65% more energy than it absorbs from the Sun. Recent calculations by Adam Burrows at the University of Arizona in the US and Tristan Guillot at the Observatoire de la Côte d'Azur in France show that the shrinkage rate for a giant planet with a given mass can be slowed considerably if it is prevented from radiating efficiently. The extreme irradiation experienced by the hot Jupiters due to their close proximity to their stars has precisely this effect, as the planet can only radiate efficiently from its dark side.

Much of the shrinkage occurs early in a planet's history when it is hot and has a large radiating surface area. The present-day radii of the hot Jupiters are therefore quite sensitive to the ages at which the planets reached their current orbits. For example, a planet that took a long time to form and spiral in will have plenty of opportunity to radiate and shrink. On the other hand, a planet that formed quickly and spiralled in rapidly would have arrived at its new orbit with a large radius, and would subsequently find it harder to cool.

Since the ages of the parent stars can generally be determined to within a billion years or so (from their luminosities, temperatures, heavy-element abundances and axial-spin rates), observational determinations of the radii of exoplanets with known masses can provide important insights into the interior compositions and histories of these planets.

Late last year, David Charbonneau, Tim Brown and others at Harvard University and the High-Altitude Observatory (HAO) in Boulder, Colorado discovered a planet in a 3.5-day orbit about the star HD 209458, which lies roughly 200 light-years from the Sun in the constellation of Pegasus. As the plane of the planetary orbit lies along our line of sight, the new

4 Sodium absorption spectra



A hot Jupiter's optical and infrared reflectivity (y-axis) depends strongly on the depth at which silicate clouds form, according to theoretical models by David Sudarsky and co-workers at the University of Arizona. If the clouds are high in the atmosphere (green), some absorption by alkali metals is seen in the optical spectrum at wavelengths less than 1 micron. As the cloud deck is pushed to lower altitudes (blue), these absorption troughs broaden until almost the entire optical spectrum is absorbed (red). Water and carbon monoxide give an even stronger "greenhouse effect" at the longer, infrared wavelengths. If the cloud deck is very low, the best place to look for reflected starlight is at the blue end of the optical spectrum.

planet passes across the face of its parent star once every orbit. The Harvard–HAO team, and others, quickly found that nearly 1.5% of the star's light was blocked each time the planet crossed in front of it. This allowed the researchers to measure for the first time the relative sizes of the planet and the star directly, and so obtain the first confirmation that the hot Jupiters are indeed gas giants. With a radius 1.35 times that of Jupiter, HD 209458's planet is over-sized for its age and mass, in agreement with recent models published by Burrows' team.

Silicate clouds and the sodium greenhouse

The progress made so far in determining the properties of these other worlds is remarkable in that it has all been achieved without actually seeing the light directly from the planets themselves. Direct detection will help astronomers to pin down the properties of these planets' atmospheres and to determine when the hot Jupiters arrived in their current, bizarre orbits.

A planet must balance the amount of radiation absorbed on its sunward hemisphere against the amount radiated back into space. As this balance determines the planetary radius, astronomers are keen to find out what fraction of the incident stellar radiation is absorbed so as to understand the sizes of these planets.

A star like our Sun or 51 Pegasi pumps out most of its power at optical and near-infrared wavelengths, so the reflectivity of a planet's atmosphere at these wavelengths plays an important role in the overall energy balance. Indeed, the large radius of the planet orbiting around HD 209458 is best explained if the optical reflectivity is low, allowing the planet to absorb a large fraction of the radiation received from its star. The other side of this particular coin is that the planet's atmosphere should then reflect relatively little starlight back into space.

Theoretical models offer several good reasons why this might be the case. A deep, cloudless atmosphere of molecular hydrogen should reflect a substantial proportion of incident radiation back into space, particularly at short wavelengths (i.e. blue light) where Rayleigh scattering is efficient. At longer wavelengths, however, the incident radiation can penetrate deeper into the atmosphere before being scattered. If the incoming photons are absorbed by other molecular or atomic species along the way, they may never re-emerge from the atmosphere. Instead, their energy is converted into heat, adding to the planet's overall thermal energy.

As on Earth, water and methane are among the molecules that absorb strongly at red and near-infrared wavelengths, thereby trapping incoming stellar radiation. However, any resemblance to the Earth's atmospheric chemistry ends there. As a rule of thumb, the temperature at the top of a planet's atmosphere varies roughly as $d^{-1/2}$, where d is the distance from the star. A hot Jupiter orbiting at 1/20th of the Earth's orbital distance around a Sun-like star should therefore be about four or five times hotter than the Earth's cosy 300 K. Temperatures ranging from 1300 to 1500 K are hot enough for substantial amounts of the alkali metals to be present in the gaseous state.

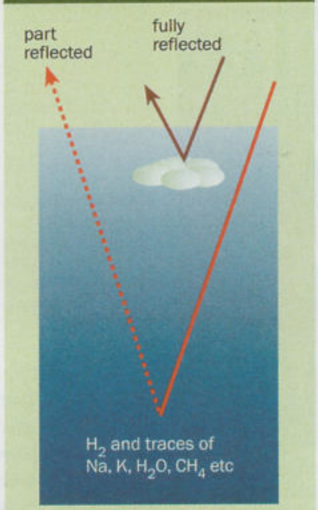
In a high-pressure atmosphere, alkali-metal atoms constantly collide with the hydrogen molecules that dominate the gas. During these collisions, the atomic energy levels are perturbed, allowing the alkali-metal atoms to absorb light at wavelengths very different from the usual narrow ranges available to isolated atoms. As a result, the absorption signature of the familiar yellow sodium "D-lines" can become so broad that it absorbs photons over almost the entire optical spectrum. This is expected to give the planets a highly effective "stealth coating", allowing very little starlight to be reflected back into space (figure 4).

The major uncertainty in this cosy picture of exoplanetary weather is the role of clouds. The pressure in any planetary atmosphere decreases with height, and so does the temperature. Clouds form if the temperature drops sharply enough with height to cross the condensation curve for any common molecule present in the atmosphere (figure 5). On Earth, the dominant cloud-forming molecule is water, and cloud systems appear brilliant white when viewed from above. Meanwhile, the temperatures in the upper atmosphere of Jupiter

and Saturn are in the range where ammonia forms clouds. These cloud decks reflect large numbers of incoming solar photons back into space, even at long wavelengths that might otherwise be absorbed by methane. As a result, Jupiter and Saturn appear white in colour, while Uranus and Neptune have deep, cloudless atmospheres in which methane absorbs most of the red light, giving them a bluish appearance.

However, the atmospheric temperatures are so high in the giant exoplanets that the dominant cloud-forming species are expected to be silicates of magnesium, such as enstatite, and perhaps even iron. Current models indicate that the reflectivity of these planets' atmospheres can increase drastically at visible wavelengths if the silicate clouds form high enough in the atmosphere so that they can scatter photons back into space before they are absorbed by sodium (figure 4). A team led by David Sudarsky of the University of Arizona in the US predicted recently that high-altitude cloud-forming conditions could be particularly favourable in the lowest mass and most strongly irradiated hot Jupiters.

5 Exoplanetary weather



Clouds form in a planetary atmosphere wherever the temperature and pressure provide the right conditions for molecules to condense from the gaseous phase to liquid droplets or solid particles. In the hot Jupiters, the most likely cloud-forming species are iron and silicates, such as enstatite. Alkali metals such as sodium and potassium are also present but remain in the gas phase. Light entering the atmosphere will be scattered efficiently back into space by the clouds. Without clouds, the light penetrates deeper into the atmosphere, encountering many more alkali-metal atoms along its path, and generally being absorbed before it can escape back into space.

tau Boo: now you see it, now you don't

Like all forms of weather prediction, exoplanetary meteorology is a complex business in which unforeseen effects due to trace species can have a disproportionately large influence on the system as a whole. Given the importance of the optical reflectivity for determining the overall energy balance of planets, these models need to be guided by direct observation.

Two groups began searching for light reflected from exoplanets about three years ago. David Charbonneau, Bob Noyes and others at Harvard used the 10 m Keck telescope in Hawaii. Meanwhile, our team at the University of St Andrews – Keith Horne, Dave James and myself – together with Alan Penny of the Rutherford Appleton Laboratory used the 4.2 m William Herschel Telescope on La Palma to search for the faint, Doppler-shifted signature of starlight reflected from the massive planet orbiting tau Bootis.

Circling its star once every 3.3 days, tau Boo's planet is the heaviest of the hot Jupiters, with a mass of at least 3.9 times – and more probably 7 or 8 times – that of Jupiter. Both teams selected tau Boo because the planet's short orbital period and predicted large radius ensure that it intercepts more light from its star than any other of the hot Jupiters. We expected that the light reflected from this planet should therefore be brighter relative to its star than any of the other exoplanets known at the time.

Each team developed sophisticated data-analysis methods to disentangle the faint signature of the reflected starlight from that of the parent star (figure 6). The spectrum of the reflected light should contain copies of the thousands of narrow absorption lines produced by heavy elements in the star's atmosphere. As the planet orbits the star, any light reflected from the planet towards the observer is Doppler-shifted by the planet's orbital motion, so we expect to see a faint echo of the star's absorption lines moving periodically back and forth with the planet's orbital speed of 150 km s^{-1}. At the same time, the strength of the reflected signature rises and falls with the changing illumination of the planet by the star. The planet is brightest when it is on the far side of the star with its illuminated hemisphere facing towards us, but invisible when it is between us and the star.

Both groups independently developed methods for subtracting out a model of the direct starlight. We then searched deep in the resulting noise for statistical evidence that the known pattern of lines was wobbling back and forth while changing in brightness at a tempo dictated by the stellar-wobble measurements made by Marcy, Butler and co-workers.

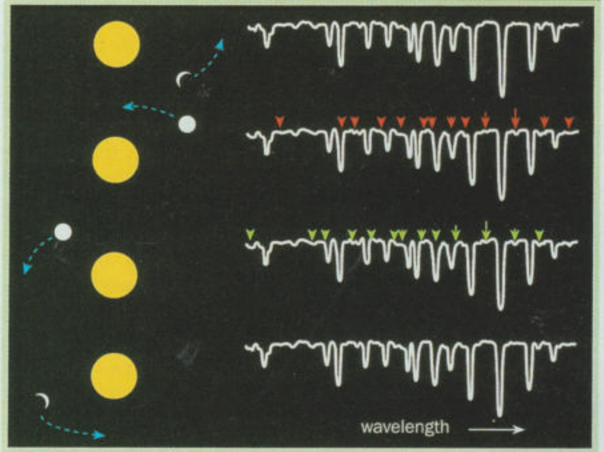
If the orbit is tilted significantly to the line of sight, the component of the planet's velocity toward the observer is lower, and the brightness variations are less pronounced. As a result, we had to search for signatures over a plausible range of orbital tilts. Charbonneau and Noyes did not detect a measurable signal during three nights of observations at Keck. Instead, they established that the planet had to be at least 10 000 times fainter than the star if its illuminated hemisphere could be viewed face on.

Our observations in April 1998, and in April and May 1999, produced a weak but plausible signal that was about 30% brighter than Charbonneau's upper limit. The result was controversial. It was hard to reconcile with the Harvard team's results, and implied that if the planet had a Jupiter-like reflectivity, its radius had to be nearly twice as large as Jupiter's. We estimated at the time that there was roughly a 5% possibility that a chance alignment of noise in our data could produce a spurious detection of this strength.

In March, April and May 2000 we observed tau Boo for a further six nights, carefully targeting those points in the orbit at which the reflected-light signature would be strongest and the absorption lines in the reflected light would be Doppler-shifted well away from the direct starlight. This strategy allowed us to probe much more sensitively for reflected light coming from a planet in the orbit suggested by our earlier measurement. However, the new observations – when combined with the 1998 and 1999 data – indicated that our earlier result had been spurious.

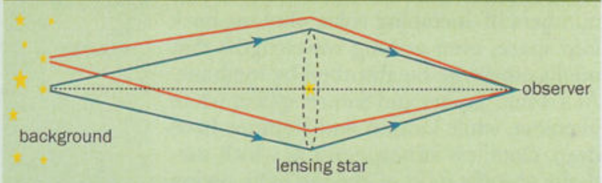
Nevertheless, our findings provided a new insight into the exoplanet's atmosphere. If the planet reflects light between 385 and 580 nm uniformly, it has to be at least 30 000 times fainter than the star. This means that if the planet's radius is 20% greater than Jupiter's – as the models of Burrows' group predict – its atmosphere must be less than 40% as reflective as Jupiter. This is well below the reflectivity predicted for a high-altitude silicate cloud deck, and suggests that tau Boo's planet may well have a deep cloud deck with an overlying stealth coating of sodium gas.

6 Reflections from another planet



Light reflected by the planet's atmosphere will contain copies of all the narrow absorption lines caused by trace elements in the star's atmosphere. The planet's orbital motion Doppler-shifts the reflected light away from the corresponding absorption lines in the direct starlight to wavelengths indicated by the red and green arrows. The length of the arrows denotes the relative strength of the shifted absorption lines. The reflected-light signal is only appreciable when the planet is on the far side of its orbit, with its illuminated face turned towards us. The reflected spectral lines then drift steadily from longer to shorter wavelengths as the planet passes behind the star. The planetary features shown in the simulated spectra here are exaggerated by a factor of about 3000.

7 Gravitational lensing



Light from a distant star in the Galactic Bulge is bent around a star in the foreground in such a way as to form two images. The summed area of these images is greater than normal, so the starlight is amplified. If the alignment is perfect, the light of the background star will pass around the lensing star to form an "Einstein ring". A planet orbiting the foreground star and located near its Einstein ring will give a second amplification event.

Next steps

Our efforts are currently devoted to the innermost of the three planets that orbit the Sun-like star upsilon Andromedae. This planet appears to be 10 times lighter than tau Boo's planet, so it has a much lower surface gravity and a more distended atmosphere. The most recent models of cloud formation for planets near their stars suggest that the silicate cloud deck may form higher in the atmospheres of planets with low surface gravities. If there is less sodium above the clouds, this planet could be very much more reflective than the planet orbiting tau Boo. We went back to the William Herschel Telescope in October and November last year to search for starlight reflected from the planet closest to upsilon Andromedae using the same techniques we developed for tau Boo, and we are currently analysing the data.

Several groups worldwide also observed HD 209458 last summer, hoping to detect the faint spectral signature of sodium in its atmosphere at the times when the planet passes between us and the star. Rather than look for light reflected from the planet's atmosphere, this method involves searching for evidence that light passing through the planet's atmosphere has been absorbed at the wavelengths of the sodium lines. HD 209458 has a similar mass and surface gravity to the planet around upsilon Andromedae that we are studying. If no one succeeds in detecting this absorption, it will mean that the amount of sodium above the cloud deck is relatively small. This would augur well for a direct detection in upsilon Andromedae, so exoplanetary astronomers await these results with keen anticipation.

Ultimately, the holy grail for planet-hunters is an Earth-like planet orbiting another star. Several programmes that have the potential to detect such planets are just beginning. The first of these involves searching for the faint dips in light that would occur as an Earth-like planet passes between us and its parent star. Earth-like planets, however, are 10 times smaller in radius – and therefore 100 times smaller in area – than Jupiter-like objects. So astronomers have to be able to detect a dip that is only one part in 10 000 of the total light from the star. This accuracy is difficult to achieve from the Earth's surface because of turbulence and the variable transparency of the Earth's atmosphere. However, several space missions have been proposed that would be capable of making such precise measurements.

The likelihood of an Earth-like planet's orbit being oriented so that we can see it cross in front of its star is very low (about 1 in 200), so these missions will have to survey hundreds or thousands of stars to have a reasonable chance of success. This is true even if half or more of all Sun-like stars possess Earth-sized planets at Earth-like distances.

In autumn last year, planet-hunters were delighted to learn that one of these missions – named Eddington in honour of the great early 20th century astrophysicist Sir Arthur Eddington – has been selected as a reserve for one of the European Space Agency's (ESA) so-called flexi-missions to fly within the next decade.

Another more bizarre search technique uses the gravitational bending of light around a star to detect planets. If we look towards the densely packed star clouds at the centre of our galaxy, we find that at any given time, one star in a million has its light amplified by this "gravitational microlensing" effect, as a star (usually a faint red dwarf) passes in front of it (figure 7). The background star appears to grow brighter then fade over a few weeks. If the star in the foreground has a Jupiter-like planet, there is about a 20% chance that a second, briefer amplification will occur as planet passes across our line of sight. That probability falls to about 2% if the planet is Earth-like.

The duration of such a secondary lensing event tells us the mass of the planet. For a Jupiter-like planet it would last about a day, whereas for an Earth-like object it would last about an hour. Several groups are already monitoring microlensing events as intensively as existing telescopes permit. At the International Astronomical Union's general assembly in Manchester last summer, Penny Sackett of the Kapteyn Institute in Groningen, the Netherlands, announced that the very fact that none of these groups has seen such an event yet suggests that less than 30% of red dwarfs possess Jupiter-like planets in Jupiter-like orbits.

Other groups are planning to build a global network of automated telescopes, capable of doing the intensive bright-

ness monitoring without the need for human intervention. Although astronomers only get one shot at each planet by this method, it could provide a good census of just how common Earth- and Jupiter-like planets may be around other stars that are less massive than the Sun.

The search for extraterrestrial life

If these relatively cheap methods produce evidence that Earth-like planets are reasonably common around solar-type stars, then the incentive to study them in detail will be overwhelming. Both NASA and ESA are looking at the possibility of using networks of infrared telescopes in space to image and obtain spectra of Earth-like planets orbiting stars up to 30 light-years away.

These two missions – known as Terrestrial Planet Finder (TPF) and Darwin – have similar aims. Both propose combining the starlight collected by four or five telescopes flying in formation about 100 metres apart to form an interference pattern. The telescopes will be positioned so that the crests and troughs of the wave trains coming from the central star via the different telescopes cancel each other out. This will allow astronomers to detect and study the light from any Earth-like planets, unobscured by the glare of the parent star.

If planets are detected, astronomers will be able to search for the thumbprints of gases like water, carbon dioxide and ozone in their infrared spectra. The presence of water would suggest a relatively benign environment for life, but finding ozone would be a clincher. Ozone – and by implication, oxygen – should not be present in a planetary atmosphere unless some mechanism is constantly renewing the supply of this highly reactive gas. On Earth, the name we give to this mechanism is "life".

But TPF and Darwin will not be cheap. They will be the most technically challenging space-science missions either agency has ever attempted. NASA and ESA will probably combine the Terrestrial Planet Finder and Darwin to keep the price within their "large missions" budget. The anticipated launch date for Darwin/TPF is around 2014, so we are uniquely privileged to be living at a time when there is a realistic prospect of seeing such a long-standing and fundamental question answered within our lifetimes.

Further reading

A Collier Cameron *et al.* 1999 Probable detection of starlight reflected from the giant planet orbiting tau Bootis *Nature* **402** 6763

G W Marcy and R P Butler 1998 Detection of extrasolar giant planets *Ann. Rev. Astr. Astrophys.* **36** 57–98

M Mayor and D Queloz 1995 A Jupiter-mass companion to a solar-type star *Nature* **378** 355

M R Zapatero Osorio *et al.* 2000 Discovery of young, isolated planetary mass objects in the sigma Orionis star cluster *Science* **290** 103

Links

Alan Penny's Darwin home page has many planet-hunting links ast.star.rl.ac.uk/darwin/

Penny Sackett's home page for the PLANET microlensing planet-search collaboration www.astro.rug.nl/~psackett/

Jean Schneider's extrasolar planets encyclopaedia www.obspm.fr/encycl/encycl.html

Andrew Collier Cameron is in the School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews, Fife KY16 9SS, UK, e-mail acc4@st-and.ac.uk

Printed and bound in the United Kingdom by Thanet Press Ltd, Union Crescent, Margate, Kent.